

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
1 March 2001 (01.03.2001)

PCT

(10) International Publication Number
WO 01/14420 A2

(51) International Patent Classification⁷: **C07K 14/00**

the University of California, 12th floor, 1111 Franklin Street, Oakland, CA 94607-5200 (US). TAMAGNONE, Luca [IT/IT]; Corso Einaudi, 43, I-10129 Torino (IT).

(21) International Application Number: **PCT/US00/23365**

(74) Agent: COX, Niki, D.; Biogen, Inc., 14 Cambridge Center, Cambridge, MA 02142 (US).

(22) International Filing Date: 25 August 2000 (25.08.2000)

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(25) Filing Language: English

(26) Publication Language: English

(84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

(30) Priority Data:
60/150,576 25 August 1999 (25.08.1999) US

Published:

— Without international search report and to be republished upon receipt of that report.

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(71) Applicants (*for all designated States except US*): UNIVERSITY OF TORINO [IT/IT]; Department of Biomedical Sciences and Human Oncology, IRCC, SP 142, I-10060 Candiolo (IT). REGENTS OF THE UNIVERSITY OF CALIFORNIA [US/US]; 12th floor, 1111 Franklin Street, Oakland, CA 94607-5200 (US).

(72) Inventors; and

(75) Inventors/Applicants (*for US only*): ARTIGIANI, Stefania [IT/IT]; Corso Brunelleschi, 121/B, I-10100 Torino (IT). COMOGLIO, Paolo, M. [IT/IT]; Strada Valsalice, 183/8, I-10100 Torino (IT). GOODMAN, Corey, S. [US/US]; Regents of the University of California, 12th floor, 1111 Franklin Street, Oakland, CA 94607-5200 (US). TESIER-LAVIGNE, Marc [US/US]; Regents of

WO 01/14420 A2

(54) Title: NOVEL PLEXINS AND USES THEREOF

(57) Abstract: The invention provides methods and compositions related to novel plexins. The polypeptides may be produced recombinantly from transformed host cells and from the disclosed plexin encoding nucleic acids or purified from human cells. The invention provides isolated plexin hybridization probes and primers capable of specifically hybridizing with the disclosed plexin genes, plexin-specific binding agents such as specific antibodies, and methods of making and using the subject compositions in diagnosis, therapy and in biopharmaceutical industry. The invention also provides novel plexin neuropilin multimeric receptor complexes for semaphorins and methods of use thereof, including but not limited to, the treatment and diagnosis of neurological disease and neuroregeneration, immune modulation, and viral and oncological diseases.

NOVEL PLEXINS AND USES THEREOF

BACKGROUND OF THE INVENTION

Field of the Invention

5 The invention relates to the identification and characterization of four novel proteins that are members of the plexin family.

Summary of the Related Art

Plexin family members encode large transmembrane proteins, whose cysteine-rich extracellular domains share regions of homology with the Scatter Factor receptors (encoded by the Met gene family). The extracellular domains of plexins also contain ~500 amino acid Semaphorin domains (see below). The highly conserved cytoplasmic moieties of plexins (approx. 600 amino acids), however, have no homology with the Met tyrosine kinase domain, nor with any other known protein. Met-like receptors and their ligands, the Scatter Factors, mediate a complex biological program including dissociation of cell-cell contacts, motility and invasion (for a review see Tamagnone, L. and Comoglio, P.M. (1997) "Control of invasive growth by hepatocyte growth factor (HGF) and related scatter factors." *Cytokine Growth Factor Rev* 8, 129-142). During embryogenesis Scatter Factor-1 and Met promote the dissociation of cell layers in the somites and drive the migration of myogenic cells to their appropriate location (Bladt, F., Riethmacher, D., Isenmann, S., Aguzzi, A., and Birchmeier, C. (1995) "Essential role for the c-met receptor in the migration of myogenic precursor cells into the limb bud." *Nature* 376, 768-771; Maina, F., Casagranda, F., Audero, E., Simeone, A., Comoglio, P.M., Klein, R.a., and Ponzetto, C. (1996) "Uncoupling of grb2 from the met receptor in vivo reveals complex roles in muscle development." *Cell* 87, 531-542).

10 Met and Scatter Factor-1 are also involved in controlling neurite outgrowth and axonal guidance (Ebens, A., Brose, K., Leonardo, E.D., Hanson, M.G.J., Bladt, F., Birchmeier, C., Barres, B.A., and Tessier-Lavigne, M. (1996) "Hepatocyte growth factor/scatter factor is an axonal chemoattractant and a neurotrophic factor for spinal motor neurons." *Neuron* 17, 1157-1172; Maina, F., Hilton, M.C., Ponzetto, C., Davies, A.M., and

15 Klein, R. (1997) "Met receptor signaling is required for sensory nerve development and HGF promotes axonal growth and survival of sensory neurons." *Genes Dev* 11, 3341-3350; Maina, F., Hilton, M.C., Andres, R., Wyatt, S., Klein, R., and Davies, A.M.

20

25

30

(1998) "Multiple roles for hepatocyte growth factor in sympathetic neuron development." *Neuron* 20, 835-846).

The first clue regarding a possible function for plexins came from the finding that a novel plexin, Vespr, interacts with the viral semaphorin A39R (Comeau, M.R., Johnson, R., DuBose, R.F., Petersen, M., Gearing, P., VandenBos, T., Park, L., Farrah, T., Buller, R.M., Cohen, J.I., Strockbine, L.D., Rauch, C., and Spriggs, M.K. (1998) "A poxvirus-encoded semaphorin induces cytokine production from monocytes and binds to a novel cellular semaphorin receptor, VESPR." *Immunity*. 8, 473-482). Semaphorins are a large family of secreted and membrane-bound molecules that are characterized by an extracellular domain containing a ~500 amino acid Semaphorin domain (Kolodkin et al. (1993) "The semaphorin genes encode a family of transmembrane and secreted growth cone guidance molecules." *Cell* 75, 1389-1399). As noted above, plexins contain a more divergent but nevertheless conserved Semaphorin domain.

Semaphorins were originally characterized in the nervous system, where they have been implicated in repulsive axon guidance (Kolodkin et al. (1993) *supra*; Luo, Y., Raible, D., and Raper, J.A. (1993) "Collapsin: a protein in brain that induces the collapse and paralysis of neuronal growth cones." *Cell* 75, 217-227; Tessier-Lavigne, M. and Goodman, C.S. (1996) "The molecular biology of axon guidance." *Science* 274, 1123-1133). More recently, semaphorins have been furthermore implicated in cardiac and skeletal development (Behar, O., Golden, J.A., Mashimo, H., Schoen, F.J., and Fishman, M.C. (1996) "Semaphorin III is needed for normal patterning and growth of nerves, bones and heart." *Nature* 383, 525-528), in the immune response (Hall, K.T., Boumsell, L., Schultze, J.L., Boussiotis, V.A., Dorfman, D.M., Cardoso, A.A., Bensussan, A., Nadler, L.M., and Freeman, G.J. (1996) "Human CD100, a novel leukocyte semaphorin that promotes B-cell aggregation and differentiation." *Proc.Natl.Acad.Sci.U.S.A.* 93, 11780-11785), in the regulation of angiogenesis (Miao, H.Q., Soker, S., Feiner, L., Alonso, J.L., Raper, J.A., and Klagsbrun, M. (1999) "Neuropilin-1 mediates collapsin-1/Semaphorin III inhibition of endothelial cell motility. Functional competition of collapsin-1 and vascular endothelial growth factor-165" [In Process Citation]. *J Cell Biol* 146, 233-242), and in tumor growth and metastasis (Christensen, C.R., Klingelhofer, J., Tarabykina, S., Hulgaard, E.F., Kramerov, D., and Lukyanidin, E. (1998) "Transcription of a novel mouse semaphorin

gene, M-semaH, correlates with the metastatic ability of mouse tumor cell lines." Cancer Res. 58, 1238-1244).

Previously identified plexins have been shown to be expressed in the developing nervous system, (i.e. Plexin-A is a receptor for class 1 semaphorins (Sema-1a and Sema-1b). Moreover, Plexin-A has been shown via genetic analysis to control motor and CNS axon guidance induced by semaphorins (Winberg, M.L., Noordermeer, J.N., Tamagnone, L., Comoglio, P.M., Spriggs, M.K., Tessier-Lavigne, M., and Goodman, C.S. (1998). Plexin A is a neuronal semaphorin receptor that controls axon guidance. Cell 95, 903-916).

Thus a need exists for discovery of other members of the plexin family of proteins.

SUMMARY OF THE INVENTION

The present invention provides four novel plexin family members.

In one aspect, the invention provides an isolated nucleic acid having at least 80% nucleic acid sequence identity to a nucleotide sequence that encodes an amino acid sequence selected from the group consisting of the amino acid sequence shown in (SEQ ID NO: 2), (SEQ ID NO: 4), (SEQ ID NO: 6) and (SEQ ID NO: 8).

In other aspects, the invention provides isolated nucleic acid having at least 80% nucleic acid sequence identity to a nucleotide sequence selected from the group consisting of the nucleotide sequence shown in (SEQ ID NO: 1), (SEQ ID NO: 3), (SEQ ID NO: 5) and (SEQ ID NO: 7).

In another aspect, the invention provides a vector comprising the nucleic acid of the above-aspects.

The invention also provides an isolated polypeptide the amino acid sequence of which comprises residues 1-18, 19-518, 451-530, 601-680, 751-830, 800-1010, or 1196-1215 of SEQ ID NO: 2; 1-23, 24-507 or 1100-1119 of SEQ ID NO: 4; or 1-42, 43-600, 541-620, 691-770, 831-910, 900-1110 or 1270-1293 of SEQ ID NO: 6; or 8-49, 154-199 or 1-199 of SEQ ID NO: 8.

In another aspect, the invention provides an isolated polypeptide having at least 80% amino acid sequence identity to an amino acid sequence selected from the group consisting of the amino acid sequence shown in (SEQ ID NO: 2), (SEQ ID NO: 4), (SEQ ID NO: 6) and (SEQ ID NO: 8)

The invention also provides a chimeric molecule comprising a polypeptide of the above aspects fused to a heterologous amino acid sequence. In one embodiment the heterologous amino acid sequence is a Fc region of an immunoglobulin.

In other aspects, the invention provides an antibody that specifically binds to the polypeptides of the above aspects.

The invention also provides a method of treating, suppressing or altering a disorder involving aberrant immune regulation involving a signaling pathway between a plexin and a neuropilin in a mammal comprising the step of administering an effective amount of an agent to said mammal capable of interfering with the association between the plexin and neuropilin. Contemplated agents include a chimeric molecule or an antibody of the above aspects.

DESCRIPTION OF THE DRAWINGS

Figure 1.

(A) Phylogenetic tree of human plexins. Known family members cluster in two major groups: plexin A and plexin-B subfamilies. (B) Structural features of plexins, Met-like receptors and semaphorins. In the extracellular moieties, yellow boxes indicate the "sema" domains and blue boxes mark the cysteine-rich MRS motifs, some of which are stippled to indicate their atypical sequence; atypical MRS are also found in the *sema domain* of semaphorins. Sequence identity among *sema* domains ranges from 15-50%, as previously described (see Winberg et al., 1998 *supra*). Potential furin-like proteolytic sites are marked by red ribbons. plexin-B1 "truncated" is the product of a splicing variant (see text). plexin-D1 and plexin-C1 (VESPR) are more distant family members, since they include atypical features in their extracellular domains. The intracellular domain of plexins (SP domain) is highly conserved through all family members, and it includes two separate regions of high homology (Maestrini, E., Tamagnone, T., Longati, P., Cremona, O., Gulisano, M., Bione, S., Tamanini, F., Neel, B.G., Toniolo, D., and Comoglio, P.M. (1996) "A family of transmembrane proteins with homology to the MET-hepatocyte growth factor receptor." Proc. Natl. Acad. Sci. USA 93, 674-678) (green oval and box). Met-like receptors are disulfide-bound heterodimers and include a cytoplasmic tyrosine kinase domain (red box). Mammalian semaphorins can be either secreted or cell surface proteins. Molecular weights of representative proteins are Plexin-A1 220 kDa, Plexin-B1 250 kDa, Plexin-C1 200

kDa, HGF-R/Met 145+45 kDa (heterodimer), Sema 4D 150 kDa (transmembrane), Sema7A approx. 100 kDa (membrane GPI-linked).

Figure 2.

(a) Cell surface semaphorins specifically bind human plexins. Micrographs of the binding assays done testing i) the extracellular domain of semaphorin CD100 fused to alkaline phosphatase (CD100-AP) on COS cells transfected with *plexin-B1 cDNA*; ii) control AP on plexin-B1; iii) CD100-AP on plexin-B2; iv) CD100-AP on the entire extracellular domain of plexin-B1; v) CD100-AP on isolated “plexin-B1 truncated” (including *sema* domain, 1° and 2° MRS); vi) CD100-AP on a “plexin-B1-Δsema” (including 2° and 3° MRS; vii) extracellular domain of semaphorin A39R fused to AP, on plexin-C1 (Vespr); viii) SemaK1-AP on plexin-C1. The final detection of the binding was done either using alkaline phosphatase substrates (i-vi) or by immunofluorescence (vii and viii). (B) Scatchard analysis and binding curve of CD100-AP to plexin-B1 ($K_D = 0.9 \text{ nM} \pm 0.15$).

Figure 3.

Plexins associate with neuropilins *via* specific extracellular domains. Western blots of immunoprecipitated samples from cells co-expressing neuropilins and plexins. Specific MoAbs were used, directed against the VSV-tag included in plexins or the myc-tag included in neuropilin-2 (Np2, 130 kDa). Np2 co-immunoprecipitates with plexins, such as plexin-A3 (220 kDa), the extracellular domain of plexinA1 (approx.160 kDa), and plexin-B1 (250 kDa) but not with the unrelated cell surface receptor DCC (170 kDa). Np2 can also associate a truncated form of the extracellular moiety of plexin-B1 (“plex-B1 trunc.”, approx. 110 kDa), containing the *sema domain*.

Figure 4.

Expression of mRNAs for plexins A1 (panel A, B), -A2 (panel C, D) and A3 (panel E, F) in the spinal cord (sc), dorsal root ganglia (d) and sympathetic ganglia (sg) of E13.5 mouse embryos. Expression of the mRNAs was detected by RNA in situ hybridization. Scale bar: 1 μm .

Figure 5.

Effect of a truncated plexin-A1 construct (lacking the cytoplasmic domain) on repulsive and attractive responses of Xenopus spinal neurons to Sema3A and netrin-1. (A-F) A control spinal neuron exposed to a gradient of Sema3A emanating from a pipette (A) is repelled away over a period of 1 hr (B). In contrast, a GFP-expressing

spinal neuron from an embryo injected with mRNA for the truncated plexin-A1 construct (C) is not affected by Sema3A (D). A similar neuron (E) shows a normal attractive response to netrin-1 (F).

(G) Cumulative distribution plot of turning angles for all the neurons studied.

5 Curves show the percent of neurons with turning angles less than the angle indicated on the abscissa, under different conditions (open circles, control neurons; black and blue circles, control neurons responding to Sema3A or netrin-1, respectively; red and green circles, responses of neurons expressing the truncated plexin-A1 construct to Sema3A and netrin-1, respectively. (H) Mean turning angle under all the conditions just
10 mentioned.

Figure 6.

Tyrosine phosphorylation of plexin-A3 and plexin-B1. (a) Anti-phosphotyrosine western blotting of immunoprecipitated $p220^{\text{plex-A3}}$ and $p250^{\text{plex-B1}}$ proteins. plexin-B1 is larger since it contains an extra sequence between the second and
15 the third MRS motif, in the extracellular domain (see Fig. 1). (b) The same immunoprecipitated samples underwent *in vitro* kinase assay in the presence of [$\gamma^{32}\text{P}$]ATP, Mg⁺⁺ and Mn⁺⁺ ions. The SDS-PAGE was treated with alkali in conditions known to eliminate the phosphate labeling of Ser/Thr residues and specifically preserving phosphotyrosines.

20 **Figure 7**

Plexin-A3 overexpression mediates cell repelling cues. (a) Epithelial kidney MDCK cells transfected to overexpress plexin-A3 (or mock transfected) were cocultured with mesenchymal KJ-29 or NIH-3T3 cells. After 16-30 hours, mixed cultures of control cells (upper panels) reached confluence and stopped growing:
25 typically the epithelial cells formed islets (circled) surrounded by a fibroblasts lawn. In contrast, MDCKs overexpressing plexin-A3 (lower panels) overwhelmed the adjacent mesenchymal cells. The latter withdrew and selectively detached from the culture dish (dying cell clusters are indicated by arrowheads), and eventually epithelial cells only survived. To allow an easier detection of the mesenchymal cells, these were labeled
30 with DiI before being plated in mixed cultures. (b) Plexin-A3 expressing cells do not induce apoptotic signal on repelled fibroblasts. Mixed cultures of NIH 3T3 and control or plexin-A3 overexpressing MDCKs were tested for the presence of TUNEL-AP positive cells. Apoptotic cells were not present in clusters of repelled cells (indicated by

arrows). The right panel shows a positive control where apoptosis was induced on the same cells by UV treatment. (c) Plexin-A3 over-expressing cells form very transient contacts with fibroblasts. Time-lapse video-microscopy of control and plexin-A3 overexpressing MDCK cells grown in presence of fibroblasts. On top, snap-shot images from the movie, taken every 50 minutes (real time). In the upper row is shown the persistent contact of a fibroblast (marked by an arrow) with an islet of control MDCK cells (marked by a star). In the lower row another fibroblast, instead, forms a transient contact with an islet of plexin-A3 transfected cells, which also, in turn, reshapes. At the bottom, the diagrams show the relative frequency of persistent, transient or very transient contacts between fibroblasts and MDCK cells.

DETAILED DESCRIPTION OF THE INVENTION

The reference works, patents, patent applications, and scientific literature, including accession numbers to GenBank database sequences, that are referred to herein establish the knowledge of those with skill in the art and are hereby incorporated by reference in their entirety to the same extent as if each was specifically and individually indicated to be incorporated by reference. Any conflict between any reference cited herein and the specific teachings of this specification shall be resolved in favor of the latter. Likewise, any conflict between an art-understood definition of a word or phrase and a definition of the word or phrase as specifically taught in this specification shall be resolved in favor of the latter.

Four novel human plexins have been identified: plexin-B2, plexin-B3, plexin-D1 and Plexin A-4. Plexin-A4 is located on human chromosome 7 and is a family member of the plexin-A subfamily which also includes plexin-A1 (alternatively named plexin-1/NOV), plexin-A2 (alternatively named plexin-2/OCT) and plexin-A3 (alternatively named plexin-2/SEX). Plexin-B2 and plexin-B3 are located on human chromosome 22 and chromosome X, respectively, and are family members of the plexin-B subfamily which also includes plexin-B1 (alternatively named SEP). Plexin-B3 maps very close to the plexin-A3 genomic locus on Xq28. Plexin-D1 is the first identified member of the plexin-D subfamily and is atypical of any of the other subfamilies. A fourth 30 subfamily of plexins, the plexin-C subfamily, is defined by VESPR (now plexin-C1).

The four novel plexins as described herein have in their extracellular domains regions of homology with two other protein families: (a) Scatter Factors Receptors, encoded by the *MET* oncogene family (Tamagnone and Comoglio, 1997 *supra*), and (b)

Semaphorins (Kolodkin et al. (1993) *supra* (Figure 1b). In particular, plexins and Met-like receptors contain short cysteine-rich motifs, termed "Met Related Sequences" (MRS), whose minimal consensus is: C-X(5-6)-C-X(2)-C-X(6-8)-C-X(2)-C-X(3-5)-C (Maestrini et al., 1996 *supra*; Tamagnone and Comoglio, 1997 *supra*); blue boxes in Fig. 1B). The proteins of the Met family contain a single MRS (in their receptor β chains), whereas in plexin family members there are two/three repeated MRS motifs. Furthermore, all plexin family members contain in their extracellular moiety a 500 amino acid region similar to the sema domain of semaphorins (Kolodkin et al. (1993) *supra*; Winberg et al., 1998 *supra*); yellow boxes in Fig. 1B. The MRS motif is proposed to function as protein-protein interaction domain.

The cytoplasmic domain of plexins contains a ~600 amino acid domain which we term the SP domain ("Sex and Plexins", marked in green in Fig. 1B) that is highly conserved within the family (57-97% similarity) and in evolution (over 50% similarity between invertebrates and humans). The SP domain does not share homology with any known protein. It includes a number of potential tyrosine phosphorylation sites, but lacks the typical motifs of catalytic tyrosine kinases. Interestingly, the predicted secondary structure of the SP domain includes long conserved alpha helices, typically found in protein-protein interaction modules. Furthermore, there are several dihydrophobic amino acid motifs (such as LL or LI), known to mediate the internalization and downregulation of transmembrane receptors (Sandoval, I.V. and Bakke O. (1994). Targeting of membrane proteins to endosomes and lysosomes. Trends in Cell Biology 4, 292-297).

The present invention also demonstrates that plexins can form complexes with neuropilins, which in turn demonstrates that a receptor for semaphorins can be hetero-oligomers of plexins and neuropilins. As demonstrated by *in situ* mRNA expression analysis, plexins and neuropilins are in fact simultaneously expressed in several neuronal populations during embryonic development. The plexin-neuropilin complex predates ligand binding, since the association is not influenced by the presence of class 3 semaphorins. That the observed plexin-neuropilin complexes are formed in *cis* is furthermore supported by the experimental conditions used (cotransfection of isolated cells with the two constructs). An interaction in *trans* might also be envisioned (considering that plexins and semaphorins share similar *sema* domains), however by

analyzing mixed cultures of cells separately transfected with plexins and neuropilins we did not isolate associated complexes (data not shown).

We observed that the main semaphorin binding domain of neuropilins (CUB domain (Giger, R.J., Urquhart, E.R., Gillespie, S.K., Levengood, D.V., Ginty, D.D., and Kolodkin, A.L. (1998) "Neuropilin-2 is a receptor for semaphorin IV: insight into the structural basis of receptor function and specificity." *Neuron* 21, 1079-1092;

Nakamura, F., Tanaka, M., Takahashi, T., Kalb, R.G., and Strittmatter, S.M. (1998) "Neuropilin-1 extracellular domains mediate semaphorin D/III-induced growth cone collapse" [In Process Citation]. *Neuron* 21, 1093-1100; Chen et al. 1998 *supra*) is not required for the interaction with plexins, as indicated by the association of the relevant

Neuropilin-2 deletion construct with plexin-B1 (not shown). A ternary complex, where neuropilins use two distinct protein modules to form a bridge between the sema domain of semaphorins and the sema domain of plexins is thus contemplated. It is further contemplated that plexins are the functional partners of neuropilins, required for transducing signals mediated by semaphorins, preferably class 3 semaphorins. For example, in flies, which lack both neuropilins and class 3 semaphorins, D Plex A appears sufficient as a functional receptor for Sema 1a, a transmembrane class 1

semaphorin (Winberg et al., 1998 *supra*). Further support that plexins are functional co-receptors for secreted semaphorins is demonstrated in an experiment that shows that a truncated plexin-A1 construct expressed in Xenopus spinal neurons abolishes repulsive responses to Sema3A without markedly affecting attractive responses to netrin-1. These results are consistent with the involvement of plexins.

The intracellular signals transduced by plexins are still largely obscure. The cytoplasmic domain of plexins is large and highly conserved within and across species and contains stretches of alpha helices, which are putative protein-protein interaction domains, and could thus mediate the association with cytosolic partners. We demonstrate herein that the cytoplasmic domain of plexins can be tyrosine phosphorylated, suggesting that, like other receptors devoid of intrinsic catalytic activity, plexins may signal by associating a tyrosine kinase (Stahl, N. and Yancopoulos, G.D. (1993). The alphas, betas, and kinases of cytokine receptor complexes. *Cell* 74, 587-590; Glass, D.J., Bowen, D.C., Stitt, T.N., Radziejewski, C., Bruno, J., Ryan, T.E., Gies, D.R., Shah, S., Mattsson, K., Burden, S.J., DiStefano, P.S.,

Valenzuela, D.M., DeChiara, T.M., and Yancopoulos, G.D. (1996). Agrin acts via a MuSK receptor complex. *Cell* 85, 513-523).

In addition, we show herein that expression of plexins, particularly plexin-A3, mediates cell-repelling cues. By time-lapse video-microscopy we observed a true 5 repelling effect on fibroblasts. Intriguingly, we observed that -upon interaction with fibroblasts- also the islets of plexin-A3 MDCKs at times reshaped. This may be explained by the existence of intra-epithelial repelling cues, balanced by the attractive forces exerted by epithelial cell junctions.

Moreover we have demonstrated that in the nervous system (i.e. *Drosophila*), that 10 defasciculating motor axons co-express both Plexin A and one of its interacting partners, the transmembrane semaphorin Sema-1a (Winberg et al., 1998 *supra*). This demonstrates that plexins act *in vivo* either as receptors or ligands for cell surface semaphorins, which in turn can transduce intracellular signals, as reported for ephrins (Holland et al., 1996 *supra*). Semaphorins, therefore, besides being pivotal in axon 15 guidance, have a general role in morphogenesis and tissue remodeling by mediating cell-repelling cues via their interactions with plexins.

Accordingly, in a first aspect, the invention provides an isolated nucleic acid molecule encoding a novel human plexin polypeptide. By "plexin polypeptide" is meant 20 a member of the plexin family comprising an amino acid sequence that shares at least 60% amino acid sequence homology with SEQ ID NOS: 2 (plexin B-2), 4 (plexin B-3), 6 (plexin D-1) or 8 (plexin A-4), preferably; at least 65% sequence homology, more preferably, at least 70% sequence homology, more preferably, at least at least 75% sequence homology, more preferably, at least 80% sequence homology, still more preferably at least 85% sequence homology, even more preferably, at least 90% sequence 25 homology, and most preferably at least 95% sequence homology with SEQ ID NOS: 2, 4, 6 or 8. Plexin polypeptides of the invention are useful for modulating cell growth (i.e. nerve) and immune regulation.

As used herein, by "modulating" is meant increasing or decreasing cell growth. By "cell growth" is meant any change in cell number or size, including, without 30 limitation, increase or decrease in cell number; increase or decrease in rate of cell division, increase or decrease in rate of cell death, and/or increase or decrease in cell size. Standard methods for measuring cell growth include standard apoptosis assays (*e.g.*, TUNEL assays, DNA fragmentation, trypan blue exclusion) and cell proliferation assays

(e.g., 3 H-thymidine incorporation). It will be appreciated that the degree of modulation of cell growth provided by a plexin polypeptide in a given assay will vary, but one of skill in the art can readily determine the statistically significant change in cell growth of a cell exposed to a plexin polypeptide.

5 By "immune regulation" is meant increasing or decreasing the biological functions of immune cells (*i.e.*, cells involved in an immune response). Immune cells include, without limitation, lymphocytes (T and B), NK cells, dendritic cells, myeloid cells (e.g., macrophages and neutrophils), and other hematopoietic cells involved in an immune response.

10 By "nucleic acid molecule" or "nucleic acid" as used herein, is meant any deoxyribonucleic acid (DNA) or ribonucleic acid (RNA), including, without limitation, complementary DNA (cDNA), genomic DNA, RNA, heteronuclear RNA (hnRNA), messenger RNA (mRNA), DNA/RNA hybrids, or synthetic nucleic acids (e.g., an oligonucleotide) comprising ribonucleic and/or deoxyribonucleic acids or synthetic variants thereof. The nucleic acid of the invention includes, without limitation, an oligonucleotide or a polynucleotide. The nucleic acid can be single stranded, or partially or completely double stranded (duplex). Duplex nucleic acids can be homoduplex or heteroduplex.

15 By "polypeptide" is meant any molecule comprising two or more amino acids joined together with a peptide bone, regardless of length or post-translational modifications (e.g., without limitation, glycosylation, lipidation, acetylation, or phosphorylation). Useful plexin polypeptides of the invention include, without limitation, the full length plexin polypeptides having the amino acid sequence of SEQ ID NOS: 2, 4, 6, 8 or 10; an extracellular domain of the polypeptide having the amino acid sequence 1 to 1199 of SEQ ID NO: 2; 1 to 1099 of SEQ ID NO: 4; 1 to 1270 of SEQ ID NO: 6 or 1 to 199 of SEQ ID NO: 8, with its associated signal peptide; or an extracellular domain of the polypeptide having the amino acid sequence 19 to 1199 of SEQ ID NO: 2; 24 to 1099 of SEQ ID NO: 4; or 43 to 1270 of SEQ ID NO: 6, lacking its associated signal peptide; an intracellular domain of the polypeptide having the amino acid sequence of SEQ ID NOS: 2, 4, 6 or 8; and polypeptides, the amino acid sequence of which comprises about residues 1-18 (putative signal sequence), 19-518 (sema domain), 451-530 (1° MRS), 601-680 (2° MRS), 751-830 (3° MRS), 800-1010 (G-P repeats) or 1196-1215 (putative transmembrane domain) of SEQ ID

NO: 2; about residues 1-23 (putative signal sequence), 24-507 518 (sema domain) or 1100-1119 (putative transmembrane domain) of SEQ ID NO: 4; or about residues 1-42 (putative signal sequence), 43-600 (sema domain), 541-620 (1° MRS), 691-770 (2° MRS), 831-910 (3° MRS), 900-1110 (G-P repeats) or 1270-1293 (putative transmembrane domain) of SEQ ID NO: 6; or about residue 8-49 (1° MRS) or 154-199 (2° MRS) of SEQ ID NO: 8.

By "isolated" is meant a compound (*e.g.*, a nucleic acid molecule or a protein) that has been separated from components (*e.g.*, nucleic acid molecules, proteins, lipids, and/or carbohydrates) which naturally accompany it. Water, buffers, and other small molecules (*e.g.*, molecules having a molecular weight of less than about 1000 daltons) may accompany an isolated compound of the invention. Preferably, an isolated compound is at least 70%, by weight, free from components which naturally accompany it. More preferably, an isolated is at least 75%, by weight, free from components which naturally accompany it; still more preferably, at least 80%, by weight, free; even more preferably, at least 85%, by weight, free; and even more preferably, at least 90%, by weight, free from components which naturally accompany it. Most preferably, a substantially purified compound is at least 95%, by weight, free from components which naturally accompany it.

Where the isolated compound is a nucleic acid molecule, the isolated nucleic acid molecule is separated from other nucleic acids (*e.g.*, genes or transcripts) or proteins which, in the naturally-occurring genome of the organism from which the nucleic acid molecule was derived, flanked the nucleic acid molecule. Isolated nucleic acid molecules therefore include, without limitation, a recombinant nucleic acid molecule incorporated into a plasmid or other vector (*e.g.*, a replication-defective virus); a recombinant nucleic acid molecule incorporated into the genome of a prokaryotic or eukaryotic organism; or a nucleic acid molecule which exists as a separate molecule independent of other nucleic acids (*e.g.*, a PCR product, a chemically synthesized nucleic acid molecule, or a nucleic acid molecule produced by restriction endonuclease digestion). Purification of a nucleic acid molecule can be accomplished and measured by any standard method including, without limitation, sequence analysis, chemical synthesis, PCR, CsCl gradient, phenol:chloroform extraction, ethanol precipitation, Southern or Northern blotting analysis followed by band extraction and purification, and

high performance liquid chromatography (HPLC; see, e.g., Fisher (1980) Laboratory Techniques in Biochemistry and Molecular Biology, Work and Burdon (eds.), Elsevier).

Thus, in one non-limiting example, to obtain an isolated nucleic acid molecule encoding a plexin polypeptide, a nucleic acid molecule is chemically synthesized on a standard oligonucleotide synthesis machine. The resulting single stranded molecule is then subjected to second strand synthesis to form a duplex DNA molecule, which is then ligated into a plasmid capable of replication in a prokaryotic or eukaryotic cell. The nucleic acid molecule is then replicated in the cell, purified (e.g., by CsCl gradient), and subjected to sequence analysis.

In certain embodiments of the first aspect of the invention, the nucleic acid molecule has a nucleic acid sequence comprising SEQ ID NOS: 1, 3, 5, 7 or 9.

Preferably, the nucleic acid molecule of the invention has not more than 500 nucleotides flanking each of the 5' and 3' ends of SEQ ID NOS: 1, 3, 5, 7 or 9. In certain embodiments, the plexin polypeptide has an amino acid sequence that comprises SEQ ID NOS: 2, 4, 6, 8 or 10. Preferably, the plexin polypeptide of the invention has not more than 50 amino acid residues flanking each of the N-terminal and C-terminal ends of SEQ ID NOS: 2, 4, 6, 8 or 10.

In certain embodiments of the first aspect of the invention, the nucleic acid molecule hybridizes under stringent conditions (as defined herein) to SEQ ID NOS: 1, 3, 5, 7 or 9.

The invention also includes nucleic acid molecules that hybridize under stringent hybridization conditions (as defined herein) to all or a portion of the nucleotide sequence represented by SEQ ID NOS: 1, 3, 5, 7 or 9 or its complement. The hybridizing portion of the hybridizing nucleic acid is at least 80%, e.g., at least 95%, or at least 98%, homologous to the sequence of a portion or all of a nucleic acid encoding a polypeptide having the amino acid sequence of SEQ ID NOS: 2, 4, 6, 8 or 10, or its complement. Hybridizing nucleic acids of the type described herein can be used, for example, as a cloning probe, a primer (e.g., a PCR primer) or a diagnostic probe.

Hybridization of the oligonucleotide probe to a nucleic acid sample typically is performed under stringent conditions. Nucleic acid duplex or hybrid stability is expressed as the melting temperature or T_m, which is the temperature at which a probe dissociates from a target DNA. This melting temperature is used to define the required stringency conditions. If sequences are to be identified that are related and substantially

identical, rather than identical, then it is useful to first establish the lowest temperature at which only homologous hybridization occurs with a particular concentration of salt (e.g., SSC or SSPE). Then, assuming that 1% mismatch results in a 1°C decrease in the T_m, the temperature of the final wash in the hybridization reaction is reduced
5 accordingly (for example, if the sequences have > 95% identity with the probe are sought, the final wash temperature is decreased 5°C). In practice, the change in the T_m can be between 0.5 C and 1.5 C per 1% mismatch. "Stringent conditions" involve hybridization at 68°C in 5x SSC/5x Denhardt's solution/1.0% SDS, and washing in 0.2xSSC/0.1% SDS at room temperature. "Moderately stringent conditions" include
10 washing in 3xSSC at 42°C. The parameters of salt concentration and temperature can be varied to achieve the optimal level of identity between the probe and the target nucleic acid. Additional guidance regarding such conditions is readily available in the art, for example, by Sambrook *et al.*, *supra*; and Ausubel *et al.*, *supra*.

Nucleic acid sequence homology (as well as amino acid sequence homology) can
15 be measured according to standard methods. Unless otherwise specified, as used herein used herein, "percent homology" of two amino acid sequences or of two nucleic acids is determined using the algorithm of Karlin and Altshul (*Proc. Natl. Acad. Sci. USA* **87**: 2264-2268, 1990), modified as in Karlin and Altshul (*Proc. Natl. Acad. Sci. USA* **90**: 5873-5877, 1993). Such an algorithm is incorporated into the NBLAST and XBLAST
20 programs of Altschul *et al.* (*J. Mol. Biol.* **215**: 403-410, 1990). BLAST nucleotide searches are performed with the NBLAST program, e (score) = 100, word length = 12, to obtain nucleotide sequences homologous to a nucleic acid molecule of the invention.
BLAST protein searches are performed with the XBLAST program, e (score) = 50, word
length = 3, to obtain amino acid sequences homologous to a reference polypeptide (e.g.,
25 SEQ ID NO: 2). To obtain gapped alignments for comparison purposes, Gapped BLAST is utilized as described in Altschul *et al.* (*Nucleic Acids Res.* **25**: 3389-3402, 1997). When utilizing BLAST and Gapped BLAST programs, the default parameters of the respective programs (e.g., XBLAST and NBLAST) are used, namely e=10; w=11 for nucleic acid; w=3 for amino acid (the Blosum 62 scoring matrix); low complexity
30 sequence filtering. The default settings of BLAST emphasize regions of local alignment to detect relationships among sequences which share only isolated regions of similarity (Altschul *et al.*, *J. Mol. Biol.* **215**: 403-410 (1990). See <http://www.ncbi.nlm.nih.gov>.

Thus, in a non-limiting example to obtain an isolated nucleic acid molecule encoding a plexin polypeptide, a nucleic acid molecule having the sequence of SEQ ID NOS: 1, 3, 5 or 7 is used to probe a cDNA library under stringent conditions according to standard techniques (see., e.g., Ausubel *et al.*, *supra*). Upon identification of a positive 5 clone (*i.e.*, a clone that hybridizes to SEQ ID NOS: 1, 3, 5 or 7 under stringent conditions), that clone is expanded and subjected to sequence analysis. A nucleic acid molecule having a nucleic acid sequence that is at least 70% identical, preferably at least 75% identical, more preferably, at least 80% identical, still more preferably at least 85% identical, even more preferably, at least 90% identical, and most preferably at least 95% 10 identical (as measured by the basic BLAST program of NCBI on default settings) to SEQ ID NOS: 1, 3, 5 or 7 is a nucleic acid molecule of the invention.

In a second aspect, the invention provides four novel isolated plexin polypeptides. “Isolated” is as defined for the first aspect of the invention. Where the isolated compound is a polypeptide, the isolated polypeptide is separated from organic molecules, 15 such as nucleic acid molecules, polypeptides, and/or carbohydrates, which, in the naturally-occurring organism from which the polypeptide was derived, accompany the polypeptide. Isolated polypeptides therefore also include a recombinant polypeptide (*e.g.*, a human polypeptide expressed in an insect cell), or a chemically synthesized polypeptide. Purification of a polypeptide can be accomplished and measured by any 20 standard method including, without limitation, chemical synthesis, recombinant polypeptide expression in prokaryotic or eukaryotic cells, affinity chromatography, Western blotting analysis, SDS-PAGE analysis, and/or HPLC.

In accordance with this aspect, the invention provides all derivatives, mutants, truncations, and/or splice variants of the four novel plexin polypeptides, so long as these 25 derivatives, mutants, truncations, and/or splice variants share at least 60% amino acid sequence homology with SEQ ID NOS: 2 ,4 ,6 or 8, preferably, at least 65% sequence homology, more preferably, at least 70% sequence homology, more preferably, at least at least 75% sequence homology, more preferably, at least 80% sequence homology, still more preferably at least 85% sequence homology, even more preferably, at least 90% 30 sequence homology, and most preferably at least 95% sequence homology with SEQ ID NOS: 2 ,4 ,6 or 8 as determined using the basic BLAST program of the National Center for Biotechnology (NCBI; National Library of Medicine, Bethesda, MD), using the

default settings defined therein using the sequence of the four novel plexin derivative, mutant, truncation and/or splice variance as the probe.

Preferred plexin polypeptide derivatives include polypeptides whose sequences differ from the sequence given in SEQ ID NOS: 2 ,4 ,6 or 8, by one or more 5 conservative amino acid substitutions, or by one or more non-conservative amino acid substitutions, deletions or insertions which do not abolish the biological activity of the plexins. Conservative amino acid substitutions typically include the substitution of one amino acid for another with similar biochemical characteristics, such as polarity, size, and/or charge. Non-limiting examples of conservative substitutions are substitutions 10 within the following groups: valine, glycine, alanine, valine, isoleucine, leucine; aspartic acid, glutamic acid, asparagine, glutamine; serine, threonine; lysine, arginine, phenylalanine, and tyrosine.

Useful methods for mutagenesis to generate plexin mutants are known in the art (see, e.g., Sambrook *et al.*, *supra*; Ausubel *et al.*, *supra*). Preferred methods for 15 mutagenesis are described in PCT Publication WO99/12965 and include PCR mutagenesis and saturation mutagenesis. A library of random amino acid sequence variants can also be generated by the synthesis of a set of degenerate oligonucleotide sequences.

In certain embodiments of the second aspect of the invention, the plexin 20 polypeptide has a sequence comprising the sequence of SEQ ID NOS: 2 ,4 ,6 or 8. In one non-limiting example, in accordance with the invention, an isolated plexin polypeptide comprising the sequence of SEQ ID NOS: 2 ,4 ,6 or 8 can chemically synthesized according to standard techniques (e.g., at a commercial peptide generating facility).

25 For example, a putative plexin polypeptide is purified and subjected to N-terminal sequencing to determine its amino acid sequence. The amino acid sequence of the polypeptide is then compared to SEQ ID NOS: 2 ,4 ,6, 8 or 10 (as measured by the basic BLAST program of NCBI on default settings). A polypeptide that shares at least 60% homology with SEQ ID NOS: 2 ,4 ,6, 8 or 10 is a plexin polypeptide of the 30 invention.

In another example, purification of a plexin polypeptide is facilitated by the addition of a tag to the polypeptide that enables purification of the tagged polypeptide. Non-limiting examples of a tag include a hemagglutinin (HA) tag, a his tag, a GST tag, a

FLAG-tag, and a myc tag. Thus, a nucleic acid molecule of the first aspect is engineered using standard molecular biology techniques to incorporate the nucleic acid sequence encoding the tag. The engineered nucleic acid molecule is then introduced and positioned for expression in an appropriate cell to produce the recombinant tagged polypeptide, which can then be readily purified by binding of the tag to its substrate. For example, the his tag binds to Ni-NTA agarose. Likewise, a GST (glutathione S-transferase) tag binds to glutathione agarose beads. Both his tag and GST tag expression and purification kits are commercially available from PharMingen (San Diego, CA). Likewise, myc-tagged plexin polypeptide are produced by cells introduced with a nucleic acid molecule encoding the tagged protein and positioned for expression in the cell.

It will be appreciated that particularly useful polypeptides of this aspect of the invention are secreted by the cell in which they are produced, thus facilitating purification of the polypeptide from the culture media in which the cells have been maintained, without requiring lysis of the cell.

In a third aspect, the invention provides a cell engineered to comprise a nucleic acid molecule encoding one of the four plexin polypeptides. By "engineered" is meant that the cell of the invention has been modified by standard molecular biology techniques. Where the cell is "engineered to comprise a nucleic acid molecule," standard molecular biology techniques have been employed to introduce the indicated nucleic acid molecule into the cell, either by transformation or transfection of the cell with a plasmid, or by infection or transduction of the cell with a recombinant virus.

The nucleic acid molecule of the first aspect of the invention is preferably subcloned into a plasmid or vector (for example, but not limited to, a vector used to generate a recombinant virus), wherein the nucleic acid molecule is positioned for expression in the plasmid or vector. The plasmid or vector is then introduced into a cell by standard techniques to produce an engineered cell in accordance with the third aspect of the invention.

In certain embodiments of the third aspect, the cell is a prokaryotic cell (*e.g.*, a bacterium). For example, *E. coli* cells (*e.g.*, DH5 α) are transformed (using, *e.g.*, electroporation) with a bacterial plasmid (*i.e.*, a plasmid containing an *E. coli* origin of replication) containing a nucleic acid molecule of the first aspect of the invention. The transformed bacteria are selected using, for example, an antibiotic-resistance encoding nucleic acid molecule (*e.g.*, ampicillin resistance) on the plasmid such that the antibiotic

resistance protein is expressed by the transformed bacteria. The transformed bacteria are then propagated, and can be cryopreserved and stored frozen in glycerol.

Those of skill in the art will appreciate that in accordance with the third aspect of the invention, a nucleic acid molecule encoding one of the four plexin polypeptides may 5 be introduced into a large variety of cells. For example, a nucleic acid molecule encoding one of the four plexin polypeptides can be introduced into prokaryotic cells (*e.g.*, bacteria), and any eukaryotic cell into which an exogenous nucleic acid molecule may be introduced. Thus, in certain embodiments of the third aspect of the invention, the cell is a eukaryotic cell. Eukaryotic cells according to this aspect of the invention that 10 comprise a nucleic acid molecule encoding one of the four plexin polypeptides include, without limitation, yeast cells, plant cells, insect cells, and mammalian cells. Within the category of mammalian cells are cells from any mammalian species (including, without limitation, mouse, hamster, monkey, human), of any lineage (including, without limitation, lymphocyte, fibroblast, stem cell), and may be an immortalized cell, or a non- 15 immortalized cell. Cells, as well as plasmids and/or vectors (*e.g.*, vectors that can be packaged to form infectious virus particles), are commercially available, for example, from the American Type Culture Collection (“ATCC”; Manassas, VA).

In certain embodiments of the third aspect of the invention, the nucleic acid molecule is positioned for expression in the cell. By “positioned for expression” is 20 meant that the nucleic acid molecule is operably linked to at least one regulatory sequence which directs the transcription and translation of the nucleic acid molecule in a cell, such that the cell engineered to comprise the nucleic acid molecule produces (*i.e.*, expresses) the protein encoded by the nucleic acid molecule. By “operably linked” is meant that the nucleic acid molecule and the regulatory sequence are connected in a such 25 way as to permit expression of the nucleic acid molecule when the nucleic acid molecule is present in a cell. Regulatory sequences include, without limitation, promoters, enhancers, IRES sequences, and polyadenylation signals. Since plexin polypeptides are involved in immune regulation and the modulation of cell growth, it may be desirable to operably link a nucleic acid molecule encoding one of the four plexin 30 polypeptides to an inducible promoter.

The four plexin polypeptides that are encoded by the nucleic acid molecules do not necessarily include the transmembrane domain of the four plexin polypeptides, and so may be produced by the cell as an intracellular polypeptide or a soluble secreted

polypeptide. For example, if the polypeptide fragment is secreted by the cell, it can be purified from the conditioned growth media of the transfected cells, without having to lyse the cells. Likewise, although a soluble intracellular polypeptide fragment is purified from only lysed cells, the fragment, being soluble, does not have to be extracted from the 5 cell membrane; thus, different lysis conditions may be used to obtain purified soluble intracellular polypeptide fragment as compared to the lysis conditions required to obtain purified full length plexin polypeptides (which has a transmembrane domain).

Protein expression systems have been established for a variety of cells and are known to those of skill in the art. Cells are also commercially available from the ATCC, 10 and a variety of protein expression system kits are commercially available from, for example, Invitrogen Corp. (Carlsbad, CA), Clontech Laboratories (Palo Alto, CA), PharMingen (San Diego, CA), Promega Corp. (Madison, WI), and Stratagene (La Jolla, CA).

For example, a nucleic acid molecule encoding one of the four plexin 15 polypeptides is operably linked to bacterial regulatory sequences (*e.g.*, T7 late promoter or bacteriophage regulatory sequences), and the resulting nucleic acid molecule is used to transform bacterial cells, where the transformed bacterial cells produce one of the four plexin polypeptides. In another example, a nucleic acid molecule encoding one of the 20 four plexin polypeptides is operably linked to baculovirus regulatory sequences in a baculovirus vector. Recombinant baculovirus are then generated and used to transduce insect cells (using, for example, the expression kit commercially available from Clontech Laboratories. The transduced insect cells comprise a nucleic acid molecule encoding one of the four plexin polypeptides positioned for expression in the insect cell, and thus produce one of the four plexin polypeptides.

Mammalian cells are widely used as protein expression systems. For example, a 25 mammalian cell may be transduced with a recombinant retrovirus or adenovirus comprising a nucleic acid molecule encoding one of the four plexin polypeptides operably linked to regulatory sequences that are either endogenous to the particular virus or exogenous to the virus (*e.g.*, a CMV promoter in a retroviral vector). The transduced 30 mammalian cell is then propagated *in vitro* in tissue culture, *in vivo* (*e.g.*, in an immunocompromised animal), and/or cryopreserved and stored frozen in DMSO.

In another example, mammalian cells are transfected with an expression plasmid comprising a nucleic acid molecule encoding one of the four plexin polypeptides

operably linked to one or more regulatory sequences on the plasmid. By "expression plasmid" is meant a plasmid in which an inserted nucleic acid molecule of interest (e.g., encoding one of the four plexin polypeptides, a plexin chimeric molecule, or tagged plexin polypeptide) is operably linked to at least one regulatory sequence such that when 5 the expression plasmid containing the inserted nucleic acid molecule of interest is introduced (e.g., by transfection) into a cell, the inserted nucleic acid molecule is positioned for expression in that cell. The nucleic acid molecule in the expression plasmid, upon being introduced into the cell, is thus positioned for expression in that cell, and enables the cell to produce one of the four plexin polypeptides encoded by the 10 nucleic acid molecule.

In one non-limiting example, a nucleic acid molecule according to the first aspect of the invention is inserted into a standard mammalian expression plasmid (e.g., pcDNA3.1 commercially available from Invitrogen Corp., Carlsbad, California), such that the inserted nucleic acid molecule encoding one of the four plexin polypeptides is 15 operably linked to the regulatory sequences in the mammalian expression plasmid. Mammalian cells are then transfected with this expression plasmid (using, e.g., CaPO₄ or DEAE-dextran). Where the expression plasmid contains an antibiotic-resistance encoding nucleic acid molecule (e.g., neomycin resistance on the pCDNA3.1 plasmid) such that the antibiotic resistance protein is expressed by the transfected cells, transfected 20 cells may be selected for the ability to grow in the presence of the antibiotic. The transfected cells may then be propagated and cryopreserved and stored in frozen in DMSO.

In a fourth aspect, the invention provides an isolated nucleic acid molecule encoding a chimeric molecule comprising at least two segments, wherein one of the 25 segments comprises one of the four plexin polypeptides. By "chimeric molecule" is meant a protein that comprises at least two segments of polypeptide joined together by any means, including, without limitation, a covalent bond (e.g., peptide bond), a non-covalent bond (e.g., ionic bond or hydrogen bond) or by a chemical crosslinker. It should be noted that one of the four plexin polypeptides that has been tagged is within the 30 definition of a chimeric molecule.

In certain embodiments of the fourth aspect of the invention, the nucleic acid molecule encoding the segment of a chimeric molecule comprising one of the four plexin

polypeptides hybridizes under stringent conditions to SEQ ID NO: 1, 3, 5 or 7.

"Stringent conditions" are as described above for the first aspect of the invention.

Standard molecular biology techniques may be employed to generate nucleic acid molecules encoding chimeric molecules according to the fourth aspect of the invention.

5 For example, a nucleic acid molecule encoding the extracellular domain of one of the four plexin polypeptides may be joined, in frame, to a nucleic acid molecule encoding the constant region of an immunoglobulin molecule (see, e.g., Chamow S.M., *Antibody Fusion Proteins*, John Wiley & Sons, New York, 1999). By "in frame" is meant that a first nucleic acid molecule is ligated to a second nucleic acid molecule such that the each 10 of the amino acid sequences of the polypeptides encoded by each of the first and the second nucleic acid molecules is not frame-shifted.

In one non-limiting example, a chimeric molecule comprising the extracellular domain of one of the four plexin polypeptides including the amino acid sequence of SEQ ID NOS: 2, 4, 6 or 8 is generated. In this example, a nucleic acid molecule encodes 15 amino acid residue number 1(19) through about amino acid residue number 1199 of SEQ ID NO: 2; amino acid residue number 1(24) through about amino acid residue number 1099 of SEQ ID NO: 4; amino acid residue number 1(43) through about amino acid residue number 1270 of SEQ ID NO: 6 and amino acid residue number 1 through about amino acid residue number 199 of SEQ ID NO: 8 with its associated signal peptide 20 (parenthesis depicts about the beginning of the amino acid sequence of the extracellular domain lacking its signal peptide). This nucleic acid molecule is fused in frame with a nucleic acid molecule encoding the constant region of an immunoglobulin, such that the chimeric molecule encoded by the resulting nucleic acid molecule generally has the following structure:

25

N-terminus	extracellular domain of SEQ ID NO: 2 with or lacking its signal peptide	amino acids from the constant region of an Ig molecule	C-terminus
N-terminus	extracellular domain of SEQ ID NO: 4 with or lacking its signal peptide	amino acids from the constant region of an Ig molecule	C-terminus
N-terminus	extracellular domain of SEQ ID NO: 6 with or lacking its signal peptide	amino acids from the constant region of an Ig molecule	C-terminus
N-terminus	extracellular domain of SEQ ID NO: 8 with or lacking its signal peptide	amino acids from the constant region of an Ig molecule	C-terminus

	peptide	
--	---------	--

The heavy chain class (*e.g.*, IgG, IgA, IgM, IgD, or IgE) can be varied in this chimeric molecule depending upon which constant region is used. Nucleic acid molecules encoding the constant region of various immunoglobulin (Ig) heavy chains are known
5 (see, *e.g.*, Zettlmeissl et al., *DNA Cell Biol.* **9**(5):347-53, 1990) Indeed, expression plasmids are available, into which the nucleic acid molecule of interest (*i.e.*, a nucleic acid molecule encoding an extracellular domain of the polypeptide of SEQ ID NO: 2; SEQ ID NO: 4; SEQ ID NO: 6; or SEQ ID NO: 8) can be inserted, and the resulting plasmid introduced into a cell to produce one of the four extracellular plexin-Ig chimeric
10 molecules (see, *e.g.*, Zettlmeissl et al., *supra*; Miller et al., *J. Exp. Med.* **178** (1): 211-222, 1993).

Any variety of chimeric molecule carrying the extracellular domains of one of the four plexin polypeptide may be generated. For example, the extracellular domain of one of the four plexin polypeptides can be myc-tagged, his-tagged, or FLAG tagged
15 according to standard molecular biology techniques.

Such extracellular proteins are particularly useful for identifying ligands to which the extracellular domain of one of the four plexin polypeptides bind. For example, extracellular plexin-D1-Ig chimera can be immobilized on a protein A-sepharose column. Molecules suspected of binding the extracellular domain of plexin-D1 are added to the
20 column, to which the molecule that binds to the extracellular domain of plexin-D1 adhere, and the non-binding molecules flow through the column. The extracellular plexin-D1-binding molecules are readily eluted, for example, by changing the pH or ion concentration of the elution buffer.

Extracellular plexin proteins are also used to identify cells expressing the ligand
25 of plexin extracellular domain on their cell surface (and thereby also identify the ligand itself). For example, cells are incubated with a FLAG-tagged plexin extracellular domain chimeric molecule. A FLAG-tagged plexin extracellular domain chimeric molecule is generated. An anti-FLAG antibody that is detectably labeled is then added to the cells. By “detectably labeled” is meant any means for marking and identifying the presence of a
30 molecule. Detectable labels include, without limitation, radioactive labels (*e.g.*, ^{32}P or ^{35}S) and fluorophore labels (*e.g.*, FITC, phycoerythrin, or rhodamine). For example, FITC-labeled anti-FLAG antibodies are commercially available from Babco, Richmond,

CA. The "stained" cells (*i.e.*, cells incubated first with the FLAG-tagged plexin extracellular domain chimeric molecule and then with the FITC-labeled anti-FLAG antibody), are then subjected to flow cytometry analysis to select those cells that are labeled with FITC, and so express a molecule that binds to the extracellular domain of one of the four plexin polypeptides. The FITC labeled cells are then further manipulated (*e.g.*, characterized to determine which cells express the plexin polypeptide ligand).

The ligand of the plexin extracellular domain is itself identified, for example, by lysing the cells, adding the lysate to one of the four plexin extracellular domain-Ig chimeric molecule columns described above, and purifying the ligand. The ligand is then sequenced by N-terminal sequencing.

In another non-limiting example, the intracellular domain of one of the four plexin polypeptides is used as "bait" in a yeast two-hybrid system to identify ligands that interact with the intracellular domain of one of the four plexins described herein. For general description of the two-hybrid system, see U.S. Patent Nos. 5,283,173; 5,468,614; 15 and 5,695,941. In this example, a nucleic acid molecule encoding from about amino acid residue number 143 through at least amino acid residue number 214 of SEQ ID NO: 2 is inserted into a standard DNA binding domain expression plasmid (*e.g.*, the GAL4 DNA binding domain plasmid in the Interactor kit commercially available from PharMingen (San Diego, CA). (It will be understood that the nucleic molecule may encode amino acid residue number 138-148 through at least amino acid residue number 214 of SEQ ID NO: 2.) A variety of cDNA libraries in transcriptional activation domain vectors are available (*e.g.*, from Clontech, Palo Alto CA). The cDNA libraries are screened employing standard methods (see, *e.g.*, the methods employed in U.S. Patent No. 5,780,262) to identify cDNA clones encoding a ligand that binds to the intracellular domain of one of the four plexin polypeptides. One preferable cDNA library screened in this example is a cDNA library generated from an immune cell (*e.g.*, a lymphocyte or NK cell).

In a fifth aspect, the invention provides a purified chimeric molecule comprising one of the four plexin polypeptides. Methods for purifying proteins are as described for the second aspect of the invention.

In a sixth aspect, the invention provides a cell engineered to comprise a nucleic acid molecule encoding a chimeric molecule comprising at least two segments, wherein one of the segments comprises one of the four plexin polypeptides. As described for the

third aspect of the invention, a nucleic acid encoding a chimeric molecule comprising one of the four plexin polypeptides may be introduced into any variety of cells. In certain embodiments, the cell is a prokaryotic cell or a eukaryotic cell. In certain embodiments, the eukaryotic cell is a yeast cell or a mammalian cell (*e.g.*, a human cell).

5 In a seventh aspect, the invention provides an isolated binding agent that specifically binds one of the four plexin polypeptides, or specifically binds a chimeric molecule comprising a segment comprising one of the four plexin polypeptides. In certain embodiments, the plexin protein has an amino acid sequence comprising SEQ ID NOS:2, 4, 6 or 8.

10 By "specifically binds" is meant a binding agent (*e.g.*, an antibody) that binds to its specific target (*e.g.*, one of the four plexin polypeptides) with greater affinity than it binds to other molecules. Preferably, where the binding agent is an antibody, the antibody preferably specifically binds to its specific target with a dissociation constant (K_D) of at least 10^{-5} M, more preferably, 10^{-6} M, even more preferably 10^{-7} M, and most 15 preferably, the binding agent specifically binds to its specific target with a K_D of at least 10^{-8} M.

10 Preferably, the binding agent of this aspect of the invention is an antibody, such as a monoclonal antibody or a polyclonal antibody, or a fragment of an antibody that specifically binds one of the four plexin polypeptides. Standard methods may be 20 employed to generate both monoclonal and polyclonal antibodies that specifically bind to one of the four plexin polypeptides of the invention. See, *e.g.*, Ausubel et al., *supra*; Coligan, J.E. et al., Current Protocols in Immunology, John Wiley & Sons, New York (1991); and Delves, P.J., Antibody Production: Essential Techniques, John Wiley & Sons, New York (1997). Briefly, the plexin polypeptides of the present invention, 25 purified according to the methods described for the second aspect of the invention, are used to immunize rabbits (*e.g.*, for polyclonal antibodies) or mice (*e.g.*, for monoclonal antibodies) to generate antibody-mediated immunity to the four plexin polypeptides used to immunize the animal. For monoclonal antibodies, antibodies can be screened by, *e.g.*, ELISA, to identify those antibodies that show the highest affinity for the immunizing 30 plexin protein or polypeptide fragment. The cloned cell producing the high affinity monoclonal antibody can then be propagated *in vitro* (where the antibody is purified from the culture supernatant) or *in vivo* (where the antibody is purified from ascites fluid), and

can also be cryopreserved and stored frozen at, e.g., -70°C in DMSO, to provide a potentially limitless supply of monoclonal antibody.

In addition to intact monoclonal and polyclonal antibodies, the invention also provides various antibody fragments, such as Fab, F(ab')₂, Fv, and sFv fragments.

5 Recombinant, chimeric, and humanized antibodies are also provided.

Recombinant "humanized antibodies" which specifically bind to one of the four plexin polypeptides can be synthesized according to methods known in the art (see, e.g., Green L.L. *et al.*, *Nature Genetics* 7: 13-21, 1994 for fully humanized antibodies expressed in transgenic animals; see also U.S. Patent Nos: 5,693,761; 5,777,085; and 10 5,585,089). Humanized antibodies are chimeras comprising mostly human IgG sequences into which at least portions of the regions responsible for specific antigen-binding (e.g., CDR's) have been inserted. Animals are immunized with the desired antigen, the corresponding antibodies are isolated, and the portion of the variable region sequences responsible for specific antigen binding are removed. The animal-derived 15 antigen binding regions are then cloned into the appropriate position of human antibody genes in which the antigen binding regions have been deleted. Humanized antibodies minimize the use of heterologous (*i.e.*, inter-species) sequences in human antibodies, and thus are less likely to elicit immune responses in the treated subject (see also, e.g., U.S. Patent No. 5,807,715).

20 Construction of different classes of recombinant antibodies can also be accomplished by making chimeric or humanized antibodies comprising nonhuman variable domains and human constant domains (CH1, CH2, CH3) isolated from different classes of immunoglobulins. For example, antibodies with increased antigen binding site valencies can be recombinantly produced by cloning the antigen binding 25 site into vectors carrying the human chain constant regions (see, e.g., Arulanandam *et al.*, *J. Exp. Med.* 177: 1439-1450, 1993).

In addition, standard recombinant DNA techniques can be used to alter the binding affinities of recombinant antibodies with their antigens by altering amino acid residues in the vicinity of the antigen binding sites. The antigen binding affinity of a 30 humanized antibody can be increased by mutagenesis based on molecular modeling (see, e.g., Queen *et al.*, *Proc. Natl. Acad. Sci.* 86: 10029-10033, 1989).

Also provided in the invention are plexin polypeptide-specific single polypeptide chain antibodies (see general methods in U.S. Patent Nos. 4,946,788 and

4,704,692); single domain antibodies (Ward, E.S. et al., *Nature* **341**: 544-546, 1989); and chimeric antibodies (U.S. Patent No. 4,816,567).

Binding agents that specifically bind the plexin polypeptides of the present invention are useful, for example, in determining expression levels of the plexin polypeptides in various tissues of the body, Western blotting analysis, and immunochromatography. Particularly, binding agents that specifically bind the plexin polypeptides are useful for binding the plexin polypeptide on a cell expressing the plexin polypeptide, thereby activating the cell.

A binding agent that specifically binds one of the four plexin polypeptides, for example, is effective as an immune modulator. Additional applications include, without limitation, an injectable formulation comprising a binding agent that specifically binds one of the four plexin polypeptides that is useful to antagonize activity in a disease involving aberrant immune regulation or a disease involving aberrant cell growth.

In an eighth aspect, the invention provides an isolated antisense oligonucleotide complementary to a portion of a nucleic acid molecule encoding one of the four plexin polypeptides. In certain embodiments, hybridization of the antisense oligonucleotide to the nucleic acid molecule inhibits transcription or translation of the nucleic acid molecule.

By two nucleic acid molecules being "complementary" to one another is meant that the first nucleic acid molecule (*e.g.*, an oligonucleotide) is able to form Watson-Crick base pair hydrogen bonds (*i.e.*, hybridize) with the second nucleic acid molecule to form a duplex. The first nucleic acid molecule is thus a "complement" of the second nucleic acid molecule.

The antisense oligonucleotides according to the invention are complementary to a region of a nucleic acid molecule (or a region at the intron/exon boundary of DNA or RNA) that encodes one of the four plexin polypeptides. Preparation of antisense oligonucleotides is well known (see, *e.g.*, Agrawal *et al.*, *Trends Biotechnol.* **10**:152-158, 1992; U.S. Patent No. 5,149,798; Agrawal *et al.*, *Proc. Natl. Acad. Sci. USA* **85**:7079-7083, 1988; Froehler, *Tetrahedron Lett.* **27**:5575-5578, 1986; and Bergot *et al.*, *J. Chromatog.* **559**:35-42, 1992).

In a ninth aspect, the invention provides a method for identifying a nucleic acid molecule encoding one of the four plexin polypeptides, comprising contacting a pool of candidate nucleic acid molecules with a nucleic acid molecule encoding one of the four

plexin polypeptides, wherein hybridization of the nucleic acid molecule encoding one of the four plexin polypeptides under stringent conditions to a candidate nucleic acid molecule identifies the candidate nucleic acid molecule as a nucleic acid molecule that encodes one of the four plexin polypeptides. According to this aspect of the invention,
5 "hybridization" and "stringent conditions" are as defined above for the first aspect of the invention. In certain embodiments, the nucleic acid molecule encoding one of the four plexin polypeptides has a nucleic acid sequence comprising SEQ ID NOS: 1, 3, 5 or 7.

It will be understood that the isolated plexin polypeptides according to the second aspect of the invention, the plexin chimeric molecules according to the fifth aspect of the invention, and binding agents that specifically bind the plexin polypeptides according to the seventh aspect of the invention, are useful as therapeutics to treat an individual suffering from, or suspected of having, a disease or disorder involving aberrant immune regulation or an individual suffering from, or suspected of having, a disease or disorder involving aberrant cell growth, particularly nerve cell growth.
10

15 By "disease or disorder involving aberrant immune regulation" is meant any disease or disorder in which an abnormal immune response is generated in response to either self or foreign antigens. Thus, this definition includes, without limitation, autoimmune diseases (*e.g.*, lupus, inflammatory bowel disease, or Diabetes Type 1) and immunosuppressive diseases (*e.g.*, multiple sclerosis or rheumatoid arthritis).

20 By "disease or disorder involving aberrant cell growth" is meant any disease or disorder in which an abnormal amount of cell growth is observed. "Cell growth" is defined above. Thus, diseases and disorders involving aberrant cell growth include hyperplasia, neoplasia, and cancer, as well as degenerative diseases, such as neurodegenerative diseases.
25

Preferable therapeutically useful plexin polypeptides are soluble polypeptides (*e.g.*, lacking the hydrophobic transmembrane domain of the plexin polypeptides), particularly soluble polypeptide fragments that are secreted by the cell in which the fragment was produced. In a preferred embodiment the soluble plexin polypeptides are selected from the group consisting of plexin-A-1 (Maestrini et al. 1996 *supra*), plexin-A-
30 2 (Maestrini et al. 1996 *supra*), plexin-A-3 (Maestrini et al. 1996 *supra*), plexin-A-4, plexin-B-1 (Maestrini et al. 1996, *supra*), plexin-B-2, plexin-B-3, plexin-C1 (Comeau et al. 1998 *supra*), plexin-D-1.

In a tenth aspect, the invention provides a method for diagnosing a disease involving aberrant immune regulation or a disease involving aberrant cell growth, comprising comparing the amino acid sequence of one of the four plexin polypeptides from an individual suspected of having the disease with the amino acid sequence of one of the four plexin polypeptides from an unaffected individual, wherein the presence of a difference between the two amino acid sequences identifies the individual suspected of having the disease as having the disease. "Disease or disorder involving aberrant immune regulation" and "disease or disorder involving aberrant cell growth" are as defined above.

By "difference" in the amino acid sequence of one of the four plexin polypeptides from an individual suspected of having the disease or disorder as compared with the amino acid sequence of one of the four plexin polypeptides from an unaffected individual, is meant any mutation that changes the amino acid sequence including substitution, deletion, or addition of one or more amino acid residues.

Thus, in one nonlimiting example, one of the four plexin polypeptides is extracted from cells of an individual suspected of having a disease involving aberrant immune regulation (*e.g.*, using an antibody according to the seventh aspect of the invention). The amino acid sequence of the plexin polypeptide is determined by N-terminal sequencing and compared to the amino acid sequence of one of the four plexin polypeptides from an unaffected individual (*i.e.*, a normal individual of the same species that does not have a disease involving aberrant immune regulation or a disease involving aberrant cell growth). If there is a difference in the two amino acid sequences, the individual suspected of having a disease involving aberrant immune regulation is identified as having a disease involving aberrant immune regulation, and may be treated accordingly.

In certain embodiments of the tenth aspect, the amino acid sequence of the plexin polypeptide from the unaffected individual comprises the sequence of SEQ ID NO: 2, 4, 6, 8 or 10.

The following examples are intended to further illustrate certain preferred embodiments of the invention and are not limiting in nature. Those skilled in the art will recognize, or be able to ascertain, using no more than routine experimentation, numerous equivalents to the specific substances and procedures described herein. Such

equivalents are considered to be within the scope of this invention, and are covered by the following claims.

Practice of the present invention will employ, unless indicated otherwise, conventional techniques of cell biology, cell culture, molecular biology, microbiology, recombinant DNA, protein chemistry, and immunology, which are within the skill of the art. Such techniques are described in the literature. See, for example, **Molecular Cloning: A Laboratory Manual**, 2nd edition. (Sambrook, Fritsch and Maniatis, eds.), Cold Spring Harbor Laboratory Press, 1989; **DNA Cloning**, Volumes I and II (D.N. Glover, ed), 1985; **Oligonucleotide Synthesis**, (M.J. Gait, ed.), 1984; U.S. Patent No. 10 4,683,195 (Mullis et al.,); **Nucleic Acid Hybridization** (B.D. Hames and S.J. Higgins, eds.), 1984; **Transcription and Translation** (B.D. Hames and S.J. Higgins, eds.), 1984; **Culture of Animal Cells** (R.I. Freshney, ed). Alan R. Liss, Inc., 1987; **Immobilized Cells and Enzymes**, IRL Press, 1986; **A Practical Guide to Molecular Cloning** (B. Perbal), 1984; **Methods in Enzymology**, Volumes 154 and 155 (Wu et al., eds), 15 Academic Press, New York; **Gene Transfer Vectors for Mammalian Cells** (J.H. Miller and M.P. Calos, eds.), 1987, Cold Spring Harbor Laboratory; **Immunochemical Methods in Cell and Molecular Biology** (Mayer and Walker, eds.), Academic Press, London, 1987; **Handbook of Experiment Immunology**, Volumes I-IV (D.M. Weir and C.C. Blackwell, eds.), 1986; **Manipulating the Mouse Embryo**, Cold Spring Harbor 20 Laboratory Press, 1986.

The following Examples are provided to illustrate the present invention, and should not be construed as limiting thereof.

EXAMPLES

Example 1

Identification and cDNA cloning of *plexins* and sequence analysis

Since the coding sequences of human *plexin-B1(SEP)*, *plexin-A2(OCT)* and *plexin-A1(NOV)* were incomplete, we obtained the missing cDNA by RT-PCR; primers were designed by homology to orthologous murine sequences and corresponding ESTs. Updated database entries are X87904, X87831 and X87832, respectively. Partial cDNA of *plexin-A4* was obtained by assembling five overlapping human ESTs (HGI THC Report: THC203425), deriving from chromosome 7 specific cDNA pools. Another EST from chr. 7 (clone 7B19F10) encodes the cytoplasmic domain of a plexin and presumably derives from the same gene as *plexin-A4*. *Plexin-B2* cDNA was amplified by RT-PCR starting from the partial cDNA sequences of clones *MM1* (Shinoura, N., Shamraj, O.I., Hugenholtz, H., Zhu, J.G., McBlack, P., Warnick, R., Tew, J.J., Wani, M.A., and Menon, A.G. (1995). Identification and partial sequence of a cDNA that is differentially expressed in human brain tumors. *Cancer Lett* 89, 215-221) and *KIAA0315* (Genbank database); the genomic locus of *SEP-B* was identified due to its 100% sequence identity with clone C22_311 from human chromosome 22. *Plexin-B3* coding sequence was identified in the genomic sequence of ALD locus on human chromosome Xq28, using the algorithms HEXON and GENIE. *Plexin-D1* was similarly found in the genomic sequence of chromosome 3 (pac pDJ70i11). The genomic sequence of *plexin-B1(SEP)*, in the region of the alternative splicing of the extracellular domain, was obtained using the following primers: sense 5'GCAGCACCTGTGCACCCACAAGGC3' and antisense: 5'TGCAGGCTGGACGGGAGGGATGAGG3'. The common donor site is CCATCAG/gtgattgt (position 2028 from ATG); the alternative splice acceptor sites are: (i) cccccttcag/AGCCC, leading to the canonical plexin-B1 sequence, and (ii) ctcccttcag/GTGAT, leading to "plexin-B1 truncated" variant. All these new sequences were analyzed using the algorithms BLAST2, NETPHOS (phosphorylation prediction sites, by Nicolaj Blom), PH-PREDICT and Consensus Protein Secondary Structure prediction at IBCP. The phylogenetic tree was generated using AllAll algorithm of the Darwin sequence analysis system (at CBRG).

Example 2

Plexin cDNA expression constructs and protein analysis

Cell transfections were carried out by Calcium phosphate and DEAE-dextran methods, using 5-10 µg of each cDNA (1-2 µg each in case of cotransfections). For transient transfections in COS and BOSC-23 cells the cDNA was cloned in pCDNA3 or 5 derived expression plasmids (Invitrogen). MDCK stable transfecants for *plexin-A3* were obtained using pCEP4 expression plasmid (Invitrogen); the selection was done in the presence of Hygromycin-B (100-200 µg/ml). Plexin-A3 positive clones of MDCK cells were isolated from two independent transfections, and showed identical biological properties. Plexin and neuropilin expression constructs included a VSV- and myc-tag at 10 the N' and C' protein termini, respectively, detected by monoclonal antibodies anti-VSV-G (cat. V-5507, Sigma) and anti-cMyc-tag (cat. OP10-100UG, Calbiochem).

"Plexin-B1 truncated" splice variant was expressed from a cDNA fragment isolated by RT-PCR and VSV-tagged at the N' terminus: the encoded amino acid sequence spans up to aa 676 (including the *sema* domain and two MRS motifs). "Plexin-B1-sema" derives from a further deletion of the plexin-B1 extracellular domain, and exclusively 15 includes the *sema* domain. "Plexin-B1-Δsema" protein mutant includes only the C' terminal half of plexin-B1 extracellular domain, starting from amino acid 606, i.e. excluding sema domain and first MRS but including second and third MRS, transmembrane and intracellular domains.

For immunoprecipitations, cells were lysed with EB buffer (20 mM Tris-HCl pH 7.4, 5 mM EDTA, 150 mM NaCl, 10% glycerol, 1% Triton X-100), in the presence of a cocktail of protease inhibitors and 1mM Na-ortovanadate. Immunoprecipitations were performed at 4°C for 4h with the appropriate antibodies; high stringency washes were performed, in the presence of 1 M LiCl.

For *in vitro* kinase assays, immunopurified proteins were incubated with kinase buffer (50 mM Hepes, 100 µM DTT, 5 mM MnCl₂, 5 mM MgCl₂) in the presence of redivue 5 µCi [γ -³²P] ATP (Amersham) for 10 minutes at 4°C in agitation. Samples were then submitted to SDS-PAGE and autoradiography, or analysed using a Phosphor-Imager system (Molecular Dynamics). Alkali treatment of the polyacrilamide gels was 30 performed with 1M KOH for two hours at 55°C.

Western blots were performed according to standard methods. Specific detection of phospho-tyrosines was done with PY20 MoAb (Trasduction laboratories). Final detection was done with ECL system (Amersham).

Example 3

Semaphorin-SEAP binding assays

Soluble forms of Semaphorin extracellular domains were expressed as chimeric molecules with placental Secreted Alkaline Phosphatase (SEAP) and harvested from
5 the conditioned media of transiently transfected COS or BOSC-23 cells. Serum-free media were concentrated over 100 times using Centricon Plus-20 filters (Millipore) with a molecular weight cutoff of 100 kDa. The AP activity of these media was assessed as described (Flanagan, J.G. and Leder, P. (1990). The kit ligand: a cell surface molecule altered in steel mutant fibroblasts. Cell 63, 185-194); the specific activity of
10 chimeric molecules was approx. 1000 U/mg. Concentrated Semaphorin-SEAP were diluted as appropriate in a Hepes buffered saline, additioned with 0.2% BSA, 0.1% NaN₃, 5 mM CaCl₂ and 1 mM MgCl₂ (HBSBA). For binding assays, COS cells transiently transfected with plexins were seeded on 48 well plates to reach confluence, and then incubated with Semaphorin-SEAP preparations (approx 1-5 nM) for 90
15 minutes at room temperature. The binding was detected as described (Flanagan and Leder, 1990). Binding experiments with plexin-C1/VESPR were as described (Comeau, M.R., Johnson, R., DuBose, R.F., Petersen, M., Gearing, P., VandenBos, T., Park, L., Farrah, T., Buller, R.M., Cohen, J.I., Strockbine, L.D., Rauch, C., and Spriggs, M.K.
20 (1998). A poxvirus-encoded semaphorin induces cytokine production from monocytes and binds to a novel cellular semaphorin receptor, VESPR. Immunity. 8, 473-482; He, Z. and Tessier-Lavigne, M. (1997). Neuropilin is a receptor for the axonal chemorepellent Semaphorin III. Cell 90, 739-751).

For *in vitro* binding assays, plexin-B1 was purified from cell extracts by immunoprecipitation with anti-VSV antibody. Extracts of mock-transfected cells were
25 used as control samples. After washing, the immunocomplexes were incubated with serial dilutions of CD100-SEAP (prepared as above) for 2 hours at 4°C, in continuous agitation. Samples were then washed 3 times with HBSBA and the bound alkaline phosphatase activity was measured by colorimetric assay using p-nitro-phenyl-phosphate, as described (Flanagan and Leder, 1990). Scatchard analysis was done using
30 Equilibrate (by GertJan C. Veenstra).

Example 4

In situ hybridization analysis

RNA *in situ* hybridization was performed essentially as described (He, Z. and Tessier-Lavigne, M. (1997). Neuropilin is a receptor for the axonal chemorepellent Semaphorin III. *Cell* 90, 739-751). Briefly cDNA fragments of plexin-A1, -A2, and -A3 were used to generate ³⁵S-labeled antisense and sense RNA probes, which were 5 used for *in situ* hybridization histochemistry of cryostat sections of rat embryos.

Example 5

Xenopus turning assay

The methods for injecting mRNA encoding various constructs, and for studying the turning responses of the neurons, are exactly as described previously (Hong, K., Hinck, L., Nishiyama, M., Poo, M.M., Tessier-Lavigne, M., and Stein, E. (1999). A ligand-gated association between cytoplasmic domains of UNC5 and DCC family receptors converts netrin-induced growth cone attraction to repulsion. *Cell* 97, 927-941); Song, H., Ming, G., He, Z., Lehmann, M., Tessier-Lavigne, M., and Poo, M. (1998). Conversion of neuronal growth cone responses from repulsion to attraction by 15 cyclic nucleotides [see comments]. *Science* 281, 1515-1518).

Example 6

Mixed-culture assays and time-lapse videomicroscopy

Mock-transfected and plexin-A3 overexpressing MDCK cells were seeded with mesenchymal cells (NIH 3T3, KJ29, D17, among others), in multiwell culture plates by 20 1:4 or 1:1 ratio. NIH and KJ-29 cells were sometimes labeled by addition of DiI (1,1'-dioctadecyl-3,3,3',3'-tetramethylindocarbocyanine perchlorate, Fluka) in the culture medium, 4 hours before harvesting for the assay; clusters of cells marked with this dye are marked in blue (in light microscopy) and emit red epifluorescence (TRITC filter). The repelling effect was observed 16-30 hours after confluence, by contrast phase 25 microscopy using Leica DM IL. The progress of the assays was also monitored by time-lapse video-microscopy (320 minutes recording were converted into 1 minute play). To determine the time-length of cell contacts, for each assay, randomly chosen fibroblasts were followed during several hours and the duration of each contact between their lamellipodia and MDCK cells was measured. The doubling time of cells and their 30 viability during the assay could also be analyzed, and no differences were observed in presence of control or plexin-A3 expressing cells. Substrate adhesion of plexin-A3 overexpressing MDCKs was analyzed by counting attached cells after 30 minutes from

seeding on micro-wells coated with fibronectin, collagen or polylysine, in the absence of calf serum: no differences versus control cells were observed.

Example 7

Apoptosis detection

5 TUNEL reaction (Boehringer detection kit) was performed on mixed cultures of MDCK and NIH3T3 cells, 24 hours after seeding in a 24-well culture plate. The labeling was converted into a colorimetric signal for analysis by light microscopy using the TUNEL-AP detection kit (Boehringer). As a positive control for the induction of apoptosis, the same cells were treated with UV-C ($50 \mu\text{J}/\text{cm}^2$) or $1\mu\text{M}$ staurosporin.

10 Although the foregoing invention has been described in some detail by way of illustration and example for purposes of clarity of understanding, it will be apparent to those skilled in the art that certain changes and modifications will be practiced. Therefore, the description and examples should not be construed as limiting the scope of the invention, which is delineated by the appended claims.

15 Example 8

Plexins are specific receptors for cell surface semaphorins in vertebrates

Plexin-C1 (VESPR) has been shown to bind the soluble viral semaphorins Sema VA and VB (Comeau et al., 1998 *supra*), and we recently found that Drosophila Plexin A (D Plex A) interacts with transmembrane Sema 1a (Winberg et al., 1998 *supra*). We 20 therefore examined in vertebrates whether the extracellular domain of several different cellular semaphorins -fused to alkaline phosphatase- could bind members of the human plexin-A, -B and -C subfamilies. Multiple secreted semaphorins of class 3 (Sema3A, Sema 3C or Sema3F; see below) did not interact with plexins-A1, -A2, -A3, -B1, B2, or -C1 (data not shown). In contrast, plexin-C1(Vespr) specifically bound Sema7A(Sema- 25 K1) (Fig. 2a), a GPI-membrane linked semaphorin (class 7). This result is not entirely unexpected, since Sema7A may represent the cellular counterpart of viral semaphorin SemaVB, previously shown to interact with this plexin (Comeau et al., 1998 *supra*). More interestingly, the class 4 transmembrane semaphorin Sema4D (CD100) did interact strongly and specifically with plexin-B1 (Fig. 2a). Thus the prototypes of two 30 distinct plexin families are the receptors for members of two distinct semaphorin subclasses. We also found that Sema7A and Sema4D do not bind to neuropilin-1 or -2 alone, nor did co-transfection of either neuropilin with plexin-B1 significantly modify

its binding efficiency (not shown). Neuropilins thus seem so far to function as receptors only for vertebrate semaphorins of class 3.

The affinity constant of Sema4D for plexin-B1 was estimated by Scatchard plot to be in the subnanomolar range ($K_D = 0.9$ nM, Fig. 2b; the estimated K_D of Sema7A for plexin-C1 is 2.1 nM, not shown). These values are consistent with those observed for semaphorins-neuropilins, and fly semaphorin1-Plexin A interactions (He and Tessier-Lavigne, 1997 *supra* Winberg et al., 1998).

We used two deletion constructs of plexin-B1 to explore the semaphorin binding sites of plexins. Neither the N-terminal half of plexin-B1 extracellular domain ("plexin-B1 truncated", see previous paragraph), nor its C-terminal half ("plexin-B1- Δ sema", see Experimental Procedures) was sufficient alone to bind CD100 (see Fig. 2a), suggesting that the binding of Sema4D depends on multiple structural determinants of the extracellular domain of plexin-B1.

Example 9

15 Plexins associate with class 3 Semaphorin receptors, Neuropilins

As outlined above, secreted semaphorins of subclass 3 are known to bind neuropilins (He and Tessier-Lavigne, 1997 *supra*; Kolodkin et al., 1997 *supra*; Chen, H., Chedotal, A., He, Z., Goodman, C.S., and Tessier-Lavigne, M. (1997) "Neuropilin-2, a novel member of the neuropilin family, is a high affinity receptor for the semaphorins Sema E and Sema IV but not Sema III." *Neuron* 19, 547-559). However, the short cytoplasmic tail of neuropilins seems to be dispensable for their biological activity (Nakamura, et al (1998) *supra*), indicating the requirement of an associated coreceptor for signal transduction. Interestingly, in *Drosophila* (where neuropilins have not been identified to date) Plexin A is sufficient to mediate the biological response to semaphorin-1 in axon guidance (Winberg et al., 1998*supra*).

In an initial set of experiments, we could not observe binding of the class 3 semaphorins Sema3A(Sema III), Sema3C(Sema E) or Sema3F(Sema IV) to plexins-A1, A2, A3, B1, B2 or C1 (not shown). To test whether plexins might be coreceptors with neuropilins for class 3 semaphorins, we set up co-precipitation experiments in COS cells to test whether neuropilins may interact with plexins. Three tested plexins (plexin-A1, -A3 and -B1) associated both with neuropilin-2 (Np2, shown in Fig. 3) and neuropilin-1 (not shown). The binding was specific, inasmuch as neither neuropilin nor any plexins coimmunoprecipitated with the netrin receptor DCC, under conditions

where DCC coimmunoprecipitated with the other netrin receptor UNC5H2 (Fig. 3 and data not shown). We observed finally that the plexin-neuropilin association is mediated by the *sema domain* of plexins, as demonstrated using either the “plexin-B1 truncated” splice variant (Figure 3) or an even shorter form of the extracellular domain (“plexin-B1-sema”, see Experimental procedures, not shown).

To further support the idea of a plexin-neuropilin multimeric receptor complex for semaphorins, we show here that plexin-A3 (e.g.) is expressed in a large number of neuronal classes, including sensory, sympathetic, motor, and olfactory bulb neurons (Figure 4 and data not shown), which are known to respond to class 3 semaphorins, and which express either neuropilin-1 or neuropilin-2 or both (Chen et al., 1997 *supra*; Feiner, L., Koppel, A.M., Kobayashi, H., and Raper, J.A. (1997). Secreted chick semaphorins bind recombinant neuropilin with similar affinities but bind different subsets of neurons *in situ*. *Neuron* 19, 539-545; He and Tessier-Lavigne, 1997 *supra*; Kolodkin et al., 1997 *supra*). Thus, plexin-A3 is a candidate for a physiological coreceptor involved in mediating class 3 semaphorin effects on these axons. Other plexins may also have a role as neuropilin coreceptors in specific cell populations, such as plexin-A2, which is expressed in a subset of sensory neurons and in dorsal horn cells, and plexin-A1, which is expressed at low levels and broadly in the spinal cord (Figure 4).

To directly test the possible involvement of plexins in class 3 semaphorin signal transduction, we studied the repulsive responses of Xenopus spinal neurons to Sema3A, which is mediated by a receptor mechanism involving neuropilin-1 (Song et al., 1998 *supra*). We asked whether these responses could be altered by expression of a presumed dominant-negative plexin-A1 construct lacking the cytoplasmic domain of the protein. Transmembrane proteins can be reliably expressed in these neurons by injecting the encoding mRNA at the developmental two cell stage, allowing the embryos to grow to tadpole stage, and then removing the spinal cord and culturing the neurons (Hong et al., 1999 *supra*). We therefore injected the mRNA encoding the truncated plexin-A1 construct, together with mRNA encoding GFP (as a reporter) and then studied the responses of spinal neurons expressing GFP that were derived from these embryos. Whereas control spinal neurons are repelled by Sema3A (Figure 5A, B and Song et al. 1998 *supra*), neurons from embryos injected with mRNA for truncated plexin-A1 did not respond with either repulsion or attraction to Sema3A (Figure 5C,

D). This blocking effect appeared to be specific, since expression of a different heterologous receptor, UNC5H2, did not impair repulsion by Sema3A (Hong et al., 1997 *supra*), and since expression of the truncated plexin construct did not block attractive responses to netrin-1 (Figure 5E, F). Figure 5G, H quantifies these effects.

5 As can be seen, the effect of Sema3A is completely abolished by the truncated plexin; although there is a slight apparent decrease in the attractive effect of netrin-1 the effect is not statistically significant.

Although we have used a truncated plexin-A1 construct, this construct may be expected to interfere with the function of various plexins, since all the plexins tested 10 (A1, A3 and B1) associated with neuropilin-1. These results support a role for one or more plexins in mediating the repulsive Sema3A signal in the Xenopus spinal neurons.

Example 10

Plexins signal via a novel type of tyrosine phosphorylated cytoplasmic domain

The sequences of plexin cytoplasmic domains are highly conserved among 15 plexins but do not match any known sequences. We found that the plexin-A3 and plexin-B1 proteins are phosphorylated on tyrosine residues when overexpressed in human kidney cells (BOSC-23), as demonstrated using anti-phosphotyrosine antibodies (Fig. 6a). Furthermore, after immunoprecipitation and in vitro kinase assays, plexin-A3 and plexin-B1 became phosphorylated (Fig. 6b). Resistance to an alkali treatment (see 20 Experimental procedures) confirmed the specific phosphorylation of tyrosine residues.

The cytoplasmic domains of several receptors, including Met proteins, become tyrosine phosphorylated owing to an intrinsic kinase activity (Ullrich, A. and Schlessinger, J. (1990) "Signal transduction by receptors with tyrosine kinase activity." Cell 61, 203-212). Since the cytoplasmic domain of plexins is not similar to any bona 25 fide or atypical tyrosine kinase, this suggests that a distinct tyrosine kinase co-immunoprecipitates in association with plexins, and is responsible for their tyrosine phosphorylation. Although some additional phosphorylated proteins can be found specifically with plexin-A3 and -B1, we have not as yet identified this associated kinase. A number of endogenously expressed tyrosine kinases, namely Met, Ron, Abl 30 and Src, were not found associated with plexin-A3 by immunoprecipitation and Western blotting (not shown). Since tyrosine phosphorylated residues often function as docking sites for intracellular signal transducers (Cantley, L.C., Auger, K.R., Carpenter, C., Duckworth, B., Graziani, A., Kapeller, R., and Soltoff, S. (1991) "Oncogenes and

signal transduction." *Cell* 64, 281-302), the fact that the cytoplasmic domains of plexins are tyrosine phosphorylated further suggests that they are part of signaling complexes.

Example 11

5 Plexin-A3 expressing cells induce repulsion of co-cultured cells

Stable transfectants expressing recombinant human plexin-A3 were successfully obtained in four different cell lines: IMR32 and AF8 (human neuroblasts), and BOSC-23 and MDCK (human and canine kidney cells, respectively). We observed modest phenotypic changes in the transfected cells, which generally become flatter and larger in 10 size. The growth rate of plexin-A3 overexpressing cells was comparable to parental lines and we did not observe differences in the ability to adhere on different substrates (data not shown).

In keeping with previous report on the related Plexin of *Xenopus laevis* (Ohta, K., Mizutani, A., Kawakami, A., Murakami, Y., Kasuya, Y., Takagi, S., Tanaka, H., 15 and Fujisawa, H. (1995). "Plexin: a novel neuronal cell surface molecule that mediates cell adhesion via a homophilic binding mechanism in the presence of calcium ions." *Neuron* 14, 1189-1199), we observed a modest increase in calcium-dependent homotypic cell aggregation of plexin-A3 transfectants (not shown). Surprisingly, we found that epithelial MDCK cells overexpressing plexin-A3 mediate strong repelling 20 cues for adjacent cells. This was observed by co-culturing mock-transfected and plexin-A3 overexpressing MDCK cells together with several non-epithelial cell lines (such as NIH3T3, Kj29, and D17; Fig. 7A). Mock MDCKs grew alongside mesenchymal cells until confluence, when both cell types stopped proliferating. In contrast, when plexin-A3-overexpressing epithelial cells were grown in the same conditions, the adjacent 25 mesenchymal cells withdrew from them, and ultimately detached from the plate.

To analyze the dynamics of this repulsion process, we monitored for 36 hours, by time-lapse video-microscopy, mixed cultures of transfected MDCK cells and fibroblasts, in a number of independent experiments. At low cell density, fibroblasts showed intrinsic motility, exploring the surface of the plate with long lamellipodia and 30 filopodia, and thus coming in contacts with a high number of stationary MDCK islets. The time-length of the contacts between fibroblasts and control MDCK cells varied from 30 minutes to several hours, lasting mostly over 100 minutes. However, when fibroblasts were cultured with MDCK cells overexpressing plexin-A3, transient

contacts were observed, often lasting less than 30 minutes (see Fig. 7C). At higher cell density, fibroblasts stopped and clustered alongside the islands of control MDCKs, whereas they kept moving in a hectic fashion between the islands of plexin-A3 transfected cells (data not shown).

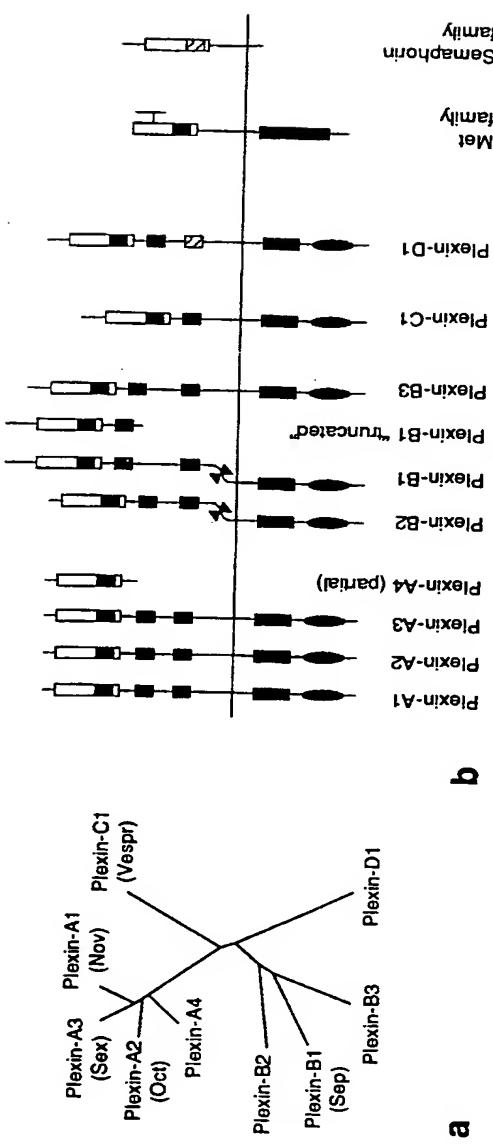
5 This cell-repelling effect is not due to the release of soluble factors, since exchanging conditioned media between mixed cultures was without effect (not shown). Moreover, the two different cell populations grew normally until they came into contact, indicating that the repelling effect requires cell-cell interaction. To rule out the possibility that plexin-A3 expressing cells generate an apoptotic signal for fibroblasts,
10 we monitored cell viability and apoptosis by TUNEL staining. As shown in Figure 7B, the clusters of repelled fibroblasts did not include apoptotic cells; furthermore, the detaching cells still excluded Trypan blue stain and were able to spread again on a new culture plate (not shown).

Taken together, these results demonstrate that in our experimental system, plexin-
15 A3 mediates cell repelling cues, presumably by interacting with surface bound ligands on opposing cells. We could not identify –so far- the specific ligand for plexin-A3, however we propose that this may be a transmembrane semaphorin. It should be noted that the intracellular domains of transmembrane semaphorins, such as Sema4D, also include tyrosine residues, which may themselves become phosphorylated and associate
20 with cytoplasmic signal transducer molecules, a property shown for ligands of the ephrin family (Holland, S.J., Gale, N.W., Mbamalu, G., Yancopoulos, G.D., Henkemeyer, M., and Pawson, T. (1996) "Bidirectional signalling through the EPH-family receptor Nuk and its transmembrane ligands." Nature 383, 722-725).

What is claimed is:

1. Isolated nucleic acid having at least 80% nucleic acid sequence identity to a
5 nucleotide sequence that encodes an amino acid sequence selected from the group consisting of the amino acid sequence shown in (SEQ ID NO: 2 (plexin B-2)), (SEQ ID NO: 4 (plexin B-3)), (SEQ ID NO: 6 (plexin D-1)) and (SEQ ID NO: 8 (plexin A-4))
2. Isolated nucleic acid having at least 80% nucleic acid sequence identity to a
10 nucleotide sequence selected from the group consisting of the nucleotide sequence shown (SEQ ID NO: 1 (plexin B-2)), (SEQ ID NO: 3 (plexin B-3)), (SEQ ID NO: 5 (plexin D-1)) and (SEQ ID NO: 7 (plexin A-4)).
3. A vector comprising the nucleic acid of any one of claims 1 or 2.
4. An isolated polypeptide having at least 80% amino acid sequence identity to
15 an amino acid sequence selected from the group consisting of the amino acid sequence shown in (SEQ ID NO: 2 (plexin B-2)), (SEQ ID NO: 4 (plexin B-3)), (SEQ ID NO: 6 (plexin D-1)) and (SEQ ID NO: 8 (plexin A-4)).
5. An isolated polypeptide having at least 80% amino-acid sequence identity
to:
 - 20 (a) the polypeptide shown in (SEQ ID NO: 2 (plexin B-2)), (SEQ ID NO: 4 (plexin B-3)), (SEQ ID NO: 6 (plexin D-1)) and (SEQ ID NO: 8 (plexin A-4)), lacking its associated signal peptide;
 - (b) an extracellular domain of the polypeptide shown in (SEQ ID NO: 2 (plexin B-2)), (SEQ ID NO: 4 (plexin B-3)), (SEQ ID NO: 6 (plexin D-1)) and (SEQ ID NO: 8 (plexin A-4)), with its associated signal peptide; or
 - (c) an extracellular domain of the polypeptide shown in (SEQ ID NO: 2 (plexin B-2)), (SEQ ID NO: 4 (plexin B-3)), (SEQ ID NO: 6 (plexin D-1)) and (SEQ ID NO: 8 (plexin A-4)), lacking its associated signal peptide.

6. A chimeric molecule comprising a polypeptide according to claim 4 or 5 fused to a heterologous amino acid sequence.
7. The chimeric molecule of claim 6, wherein the heterologous amino acid sequence is a Fc region of an immunoglobulin.
- 5 8. An antibody that specifically binds to a polypeptide according to claim 4 or 5.
9. The antibody according to claim 8, wherein the antibody is a monoclonal, a humanized antibody or a single-chain antibody.
- 10 10. A method of suppressing or altering aberrant cell growth involving a signaling pathway between a plexin and a neuropilin in a mammal comprising the step of administering an effective amount of an agent to said mammal capable of interfering with the association between the plexin and neuropilin
11. A method of treating, suppressing or altering a disorder involving aberrant immune regulation involving a signaling pathway between a plexin and a neuropilin in a mammal comprising the step of administering an effective amount of an agent to said mammal capable of interfering with the association between the plexin and neuropilin.
- 15 12. The method according to claim 10 or 11 wherein said agent is a chimeric molecule according to claim 6 or 7.
- 20 13. The method according to claim 10 or 11, wherein said agent is an antibody according to claim 8 or 9.
- 25 14. A method of diagnosing or screening for tumors in a subject characterized by the expression profiles of the polypeptides according to claim 4 or 5 wherein the expression profile of the polypeptides is different in a non-tumor sample as compared to the expression profile of the polypeptides in a tumor sample.

Figure 1 (Tammagno et al.)

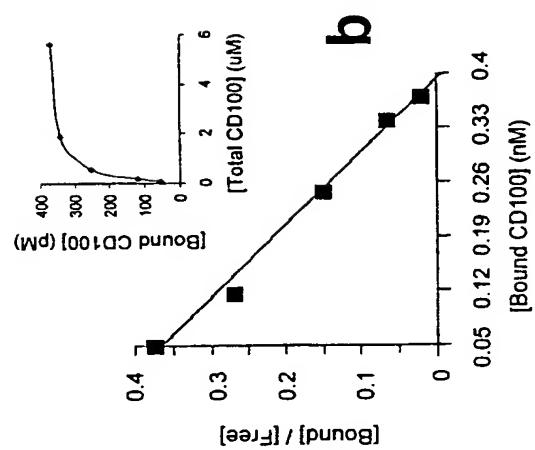
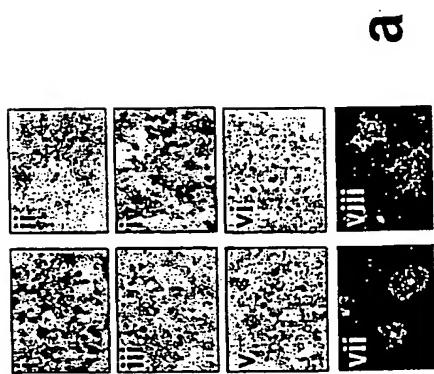
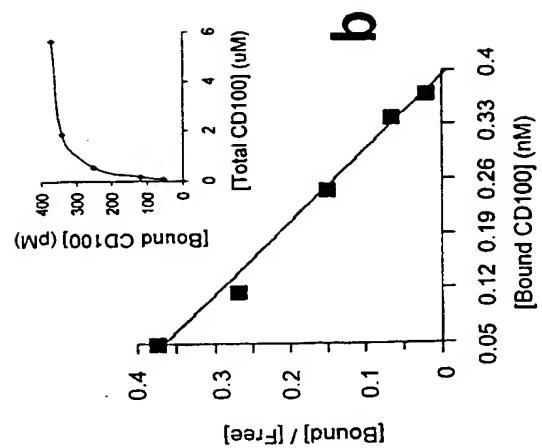
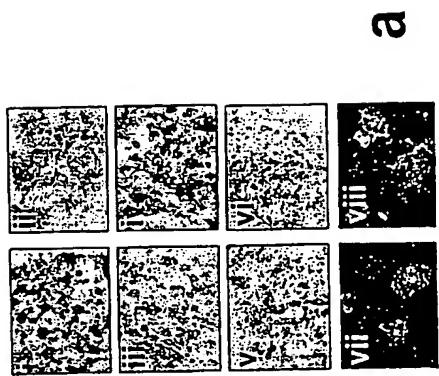


FIG. 2

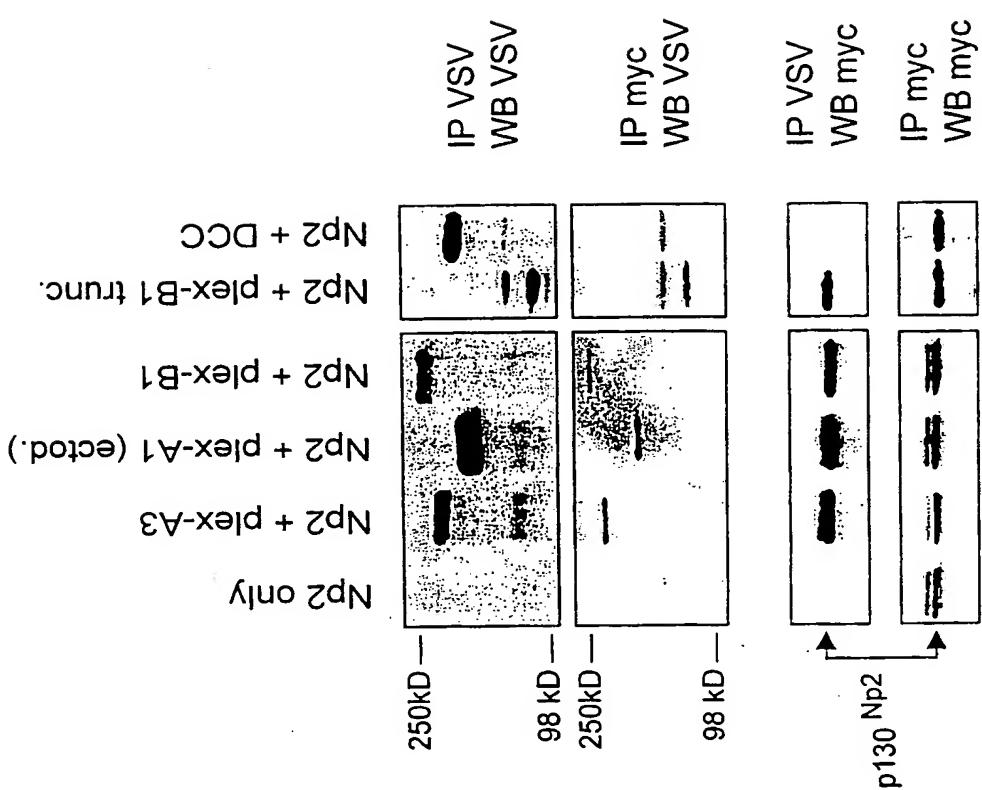


Fig. 3

Plexin-A1 Plexin-A2 Plexin-A3

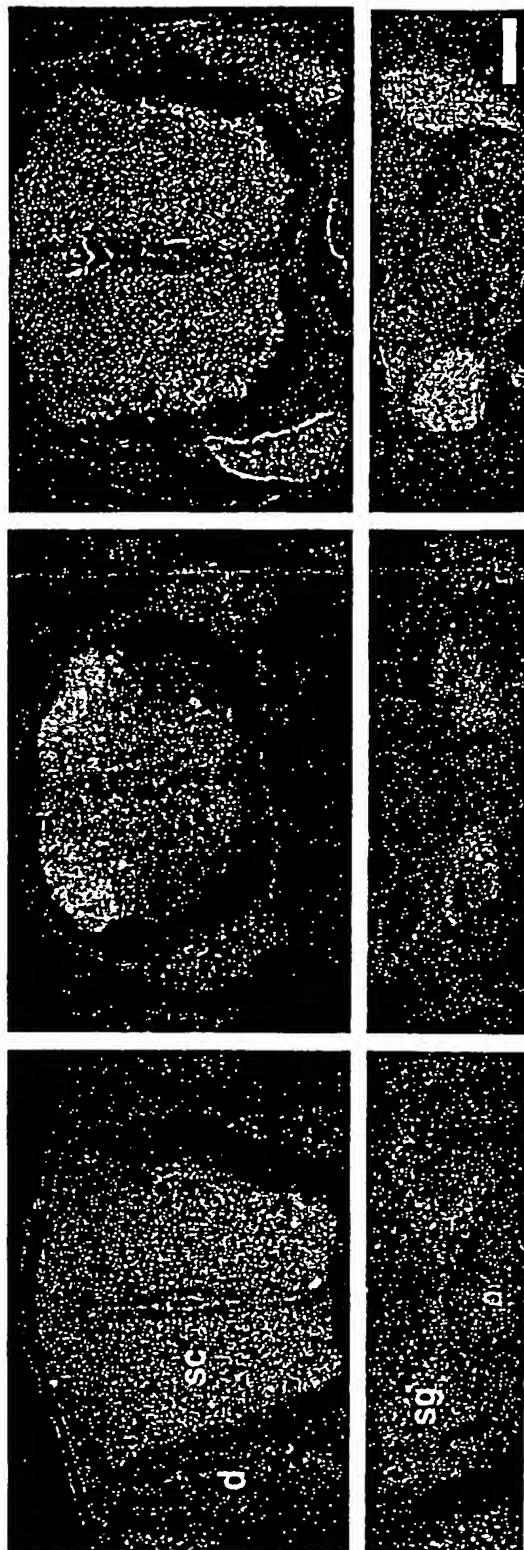
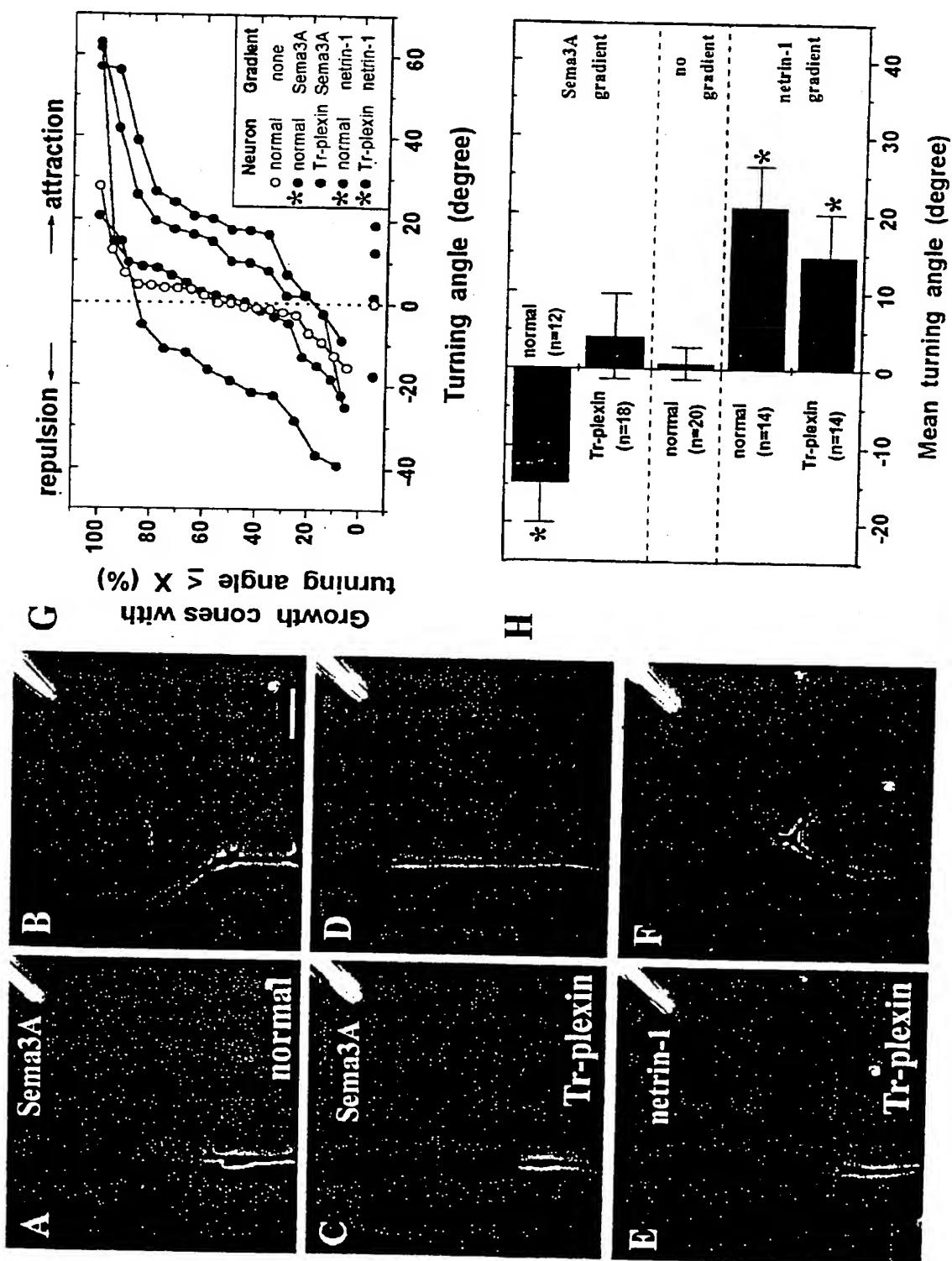


FIG. 4



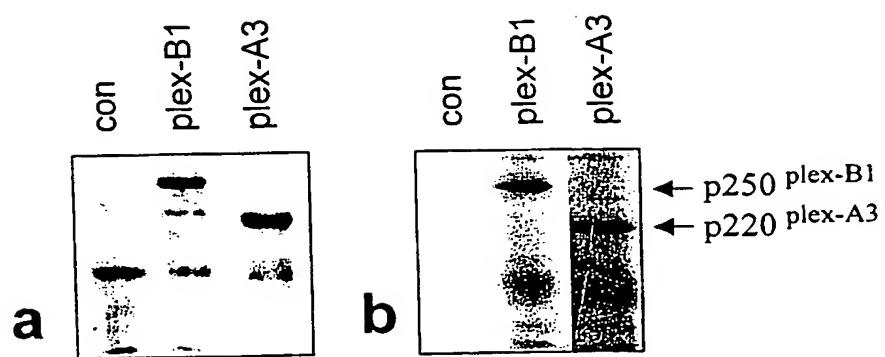
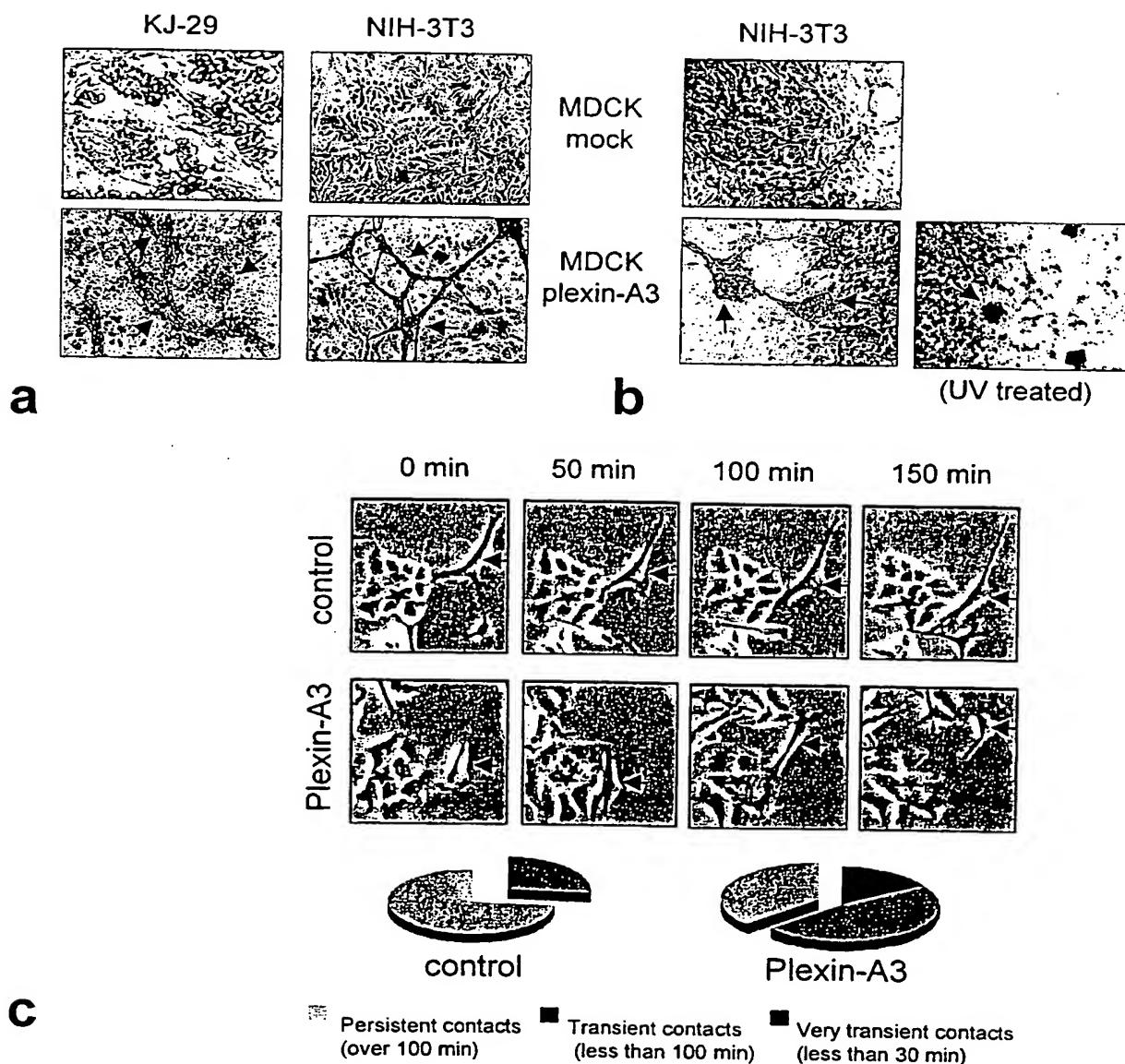


Fig 6

7/7
FIG. 7



SEQUENCE LISTING

<110> University of Torino
 <120> Novel Plexins and Uses Thereof
 <130> A077PCT
 <140> Not assigned yet
 <141> 2000-08-25
 <150> 60/150576
 <151> 1999-08-25
 <160> 12
 <170> FastSEQ for Windows Version 4.0

<210> 1
 <211> 6252
 <212> DNA
 <213> HOMO SAPIEN
 <400> 1

gcggggggca	atggcactgc	agctctgggc	cctgaccctg	ctggggcctgc	tgggcgcagg	60
tgccagcctg	aggccccgca	agctggactt	cttccgcagc	gagaaaagagc	tgaaccacct	120
ggctgtggat	gaggcctcag	gcgtgggtta	cctggggcg	gtgaatgccc	tctaccagct	180
ggatgcgaag	ctgcagctgg	agcagcagggt	ggccacgggc	ccggccctgg	acaacaagaa	240
gtgcacgccc	cccatecgagg	ccagccagggt	ccatgaggct	gagatgactg	acaatgtcaa	300
ccagctgtcg	ctgtcgacc	ctcccgaggaa	gcccctgg	gagtgcggca	gcctttcaa	360
gggcattctgc	gctctgcgcg	ccctgagcaa	catctccctc	ccgctgtttct	acgaggacgg	420
cagcggggag	aagtctttcg	tggccagcaa	tgatgagggc	gtggccacag	tggggctgg	480
gagctccacg	ggtcctgtgt	gtgaccgcgt	gctgtttgt	ggcaaaggca	atggggccaca	540
cgacaacggc	atcatcgta	gcactcggct	gttggaccgg	actgacagca	gggaggcctt	600
tgaaggctac	acggaccacg	ccacctacaa	ggccggctac	ctgtccaccca	acacacagca	660
gttcgtggcg	gccttcgagg	acggccctta	cgttttttt	gttctcaacc	acgaggacaa	720
gcacccggcc	cggAACCGCA	cgctgtggc	acgcatgtgc	agagaagaccc	ccaactacta	780
ctcc tacctg	gagatggacc	tgcaatgtgg	ggacccggac	atccacacggc	ctgcctttgg	840
cacctgcctg	gccgcctccg	tggctgcgcc	tggctctggc	agggtgctat	atgctgtctt	900
cagcagagac	agccggagca	gtggggggcc	cggtgcgggc	ctctgcctgt	tcccgcgttga	960
caaggtgcac	gccaagatgg	aggccaaccg	caacgcctgt	tacacaggca	cccgggaggc	1020
ccgtgacatc	ttctacaagc	ccttccacgg	cgatattccag	tgccggcggcc	acgcggccgg	1080
ctccacgcaag	agtttccat	gtggctcgga	gcacccgttgc	tacccgttgc	gcagccgcga	1140
cgggctcaga	ggcacagccg	tgtcgacgc	tggaggcctg	aacctcacgg	ccgtgacgg	1200
cggcccccgg	aacaaccaca	cttgtgtttt	tctgggcacc	tctgtatggcc	ggatcttcaa	1260
ggtgtacctc	accccagatg	gcaccccttc	agagtacggac	tctatcccttgc	tggagataaaa	1320
caagagagtc	aaggcgaccc	tggtaatgtc	tggagacctg	ggcagccctgt	acgcacatgac	1380
ccaggacaag	gtgttccggc	tgccggtgca	ggagtgcctg	agctaccgg	cctgcaccca	1440
gtgcgcgcac	tcccaggacc	cctactgcgg	ctgggtgcgtc	gtcgaggggac	gatgcacccg	1500
gaaggcccgag	tgtccgcggg	ccgaggaggac	cagccactgg	ctgtggagcc	gaagcaagtc	1560
ctgcgtggcc	gtcaccacgc	cccaaggccaca	gaacatgagc	cgccggggccc	agggggaggt	1620
gcagctgacc	gtcagccccc	tcctgcctt	gagcgaggag	gacgagttgc	tgtgccttt	1680
tggggagtgc	ccgcccacacc	ccggcccgct	ggagggcgag	gccgtcatct	gcaactcccc	1740

aaggcagcatc	cccgtcacac	cggccaggcca	ggaccacgtg	gccgtgacca	tccagctct	1800
ccttagacga	ggcaacatct	tcctcacgtc	ctaccagtac	cccttctacg	actggcgcca	1860
ggccatgagc	ctggaggaga	acctgccgtg	catctccctgc	gtgagcaacc	gctggacctg	1920
ccagtgggac	ctgcgctacc	acgagtgccg	ggaggcttcg	ccaaaccctg	aggacggcat	1980
cgtccgtgcc	cacatggagg	acagctgtcc	ccagttcctg	gaccccgacc	ccctggtgat	2040
ccccatgaac	cacgagacag	atgtgaactt	ccagggcaag	aacctggaca	ccgtgaaggg	2100
ttctccctg	cacgtgggca	gtgacttgct	caagtctatg	gagccggtgta	ccatgcagga	2160
atctgggacc	ttgccttgc	ggaaaaaaa	gatgtttccac	gatgccaacg	agacgctgcc	2220
cctgcaccc	tacgtcaagt	cttacggcaa	aatatcgac	agcaagctcc	atgtgaccct	2280
ctacaactgc	tcctttggcc	gcagcgactg	cagcctgtgc	cgggccgcta	accccgacta	2340
caggttgtcg	tggtgccgggg	gccaagcag	gtgcgtgtat	gaggccctgt	gcaacaccac	2400
ctccgagtgc	ccggccggcc	tcatcaccag	gatccagcct	gagacgggccc	ccctgggtgg	2460
gggcattccgc	atcaccatcc	tgggtccaa	tttgggcgtc	caaggagggg	acatccagag	2520
gatctctgtg	gcccggccgg	actgtctcc	tcagccggaa	cgtaactccg	tgtccacccg	2580
gatccgttgt	gtgatcgagg	ctgcccggac	gcctttcag	gggggtgtcg	aggtggacgt	2640
cttcggaaaa	ctggggccgtt	cgccctccaa	tgtccagttc	accttccaa	agcccaagcc	2700
tctactgtg	gagccgcagc	agggaccgca	ggcggccggc	accacactga	ccatccacgg	2760
caccaccc	gacacgggct	cccaggagga	cgtgcgggtg	accctcaacg	gcgtcccg	2820
taaagtgcac	aagtttgggg	cgca	gtgtgtca	ggcccccagg	cgacacgggg	2880
ccagatgctt	ctggaggtct	cctacggggg	gtccccctgt	ccaaacccccc	gcatcttctt	2940
cacccaccgc	aaaaaccccg	tactgcgac	cttcgagccg	ctacgaagct	ttgccagtgg	3000
tggcccgage	atcaacgtca	cgggtcagg	tttcagcctg	atccagaggt	ttgccatgtt	3060
ggtcatcgcg	gagccccc	atgtctggca	gcccggccgg	gaggctgaat	ccctgcagcc	3120
catgacgggt	gtgggtacag	actacgttt	ccacaatgc	accaggcgtc	tcttcgttc	3180
ccccggctgt	cctgaggagc	cagaggccta	caacctcag	gtgctgatcg	agatggacgg	3240
gcaccgtggc	ctgctcagaa	cagaggccgg	gcccgtcag	tacgtgcctg	acccacac	3300
tgagaacttc	acaggtggcg	tcaagaagca	ggtcaacaag	ctcatccacg	cccggggcac	3360
caatctgaac	aaggcgatga	cgctgcagga	gcccggaggcc	ttcgtgggtg	ccgagcgtcg	3420
caccatgaag	acgctgacgg	agaccgac	gtactgtgg	ccccccgggg	tgcagcccc	3480
gccccggcg	cgccagaaac	gagacaccac	acacaaac	cccggatc	ttgtgaagtt	3540
cggctctcg	gagtgggtgc	tggcccggt	ggagtagcag	acacgggtga	gcgacgtg	3600
gctcagcc	atcttggcc	tggtcatcg	gcccattgg	gtcgtcatcg	cggtgtctgt	3660
ctactgtac	tggaggaaga	gccagcaggc	cgaacgag	tatgagaaga	tcaagtccca	3720
gctggagggc	ctggagggaga	gctgtcgaa	ccgctgca	aaggaattca	cagacctgat	3780
gatcgagatg	gaggaccaga	ccaacgac	gcacgaggcc	ggcatcccc	tgctggacta	3840
caagacctac	accgaccgc	tcttcttcc	gcccattca	gacggcgaca	aggacgtat	3900
gatcaccggc	aagctggaca	tccttgc	gcccggccgg	gtgggtggag	aggccctcta	3960
ccagttctcc	aacctgtca	acagca	tttccatc	aatttcatc	acacccctgg	4020
gaaccagcgg	gagttctcg	cccgcccaa	ggtctactt	gctccctgc	tgacgggtgg	4080
gctgcacggg	aaactggagt	actacacgg	catcatgc	acgtcttcc	tggagctcct	4140
ggagcgtac	gtgggtggca	agaacccaa	gctgtatgt	cgccaggctg	agactgtgg	4200
ggagacgtat	ctgtccaa	ggatgtccat	ctgcctgtac	cagtacctc	aggacagtgc	4260
cggggagccc	ctgtaca	tcttcaaggc	catcaa	cagggtggaaa	agggcccccgt	4320
ggatgcgtt	cagaagaagg	ccaa	tctcaac	acggggctgc	tggggatg	4380
tggtggatgt	gcacccctga	cgggtgag	gtatgtcg	gacgggggg	tggacggccat	4440
cccggtgaag	gtccctcaact	gtgacaccat	ctcccgagg	aaggagaaga	tcattgacca	4500
gggtgtaccgt	gggcagccct	gtccctgt	gcccaggc	gacagcgtgg	tcctggagtg	4560
gggtccgggc	tccacagc	agatcctgtc	ggacctgg	ctgacgtc	agcgggaggg	4620
cccggtgaa	cgcgtcaaca	cccttatg	ctacaat	cgggatgg	ccaccctcat	4680
cctgtccaa	gtgggggtct	cccagcagc	ggaggac	cacggacc	tgcctggg	4740
gcccattgc	ctccgtgg	aggagaacc	gtgtgg	ctgggtcg	cgaccgac	4800
ggggacgag	ggcaagtc	agagaggc	cgtgaa	aaggagc	cgaaaggccat	4860
caccgagatc	tacctgac	ggctgtctc	agtcaagg	acactgc	tgacttgg	4920
caacttcttc	cagagcgtc	tggccctgg	gcacgcgg	ccacctgc	tgcaactt	4980
cttcgacttc	ctggacgagc	aggcagagaa	gcacaacatc	caggatga	acaccatcca	5040

catctggaag acgaacagct taccgctccg	gttctgggtg aacatcctca agaacc cccca	5100
cttcatcttt gacgtgcatt tccacgagg tgggacgccc tcgctgtcag tcatcg cgc a	5160	
gacccatg gatgcctgca cgccgacgg a cataagctg agccgcgatt ctcccagcaa	5220	
caagctgctg tacgccaagg agatctccac ctacaagaag atggtgagg attactacaa	5280	
ggggatccgg cagatggtc aggtcagcga ccaggacatg aacacacacc tggcagagat	5340	
ttccccggcg cacacggact ccttgaacac ctcgtggca ctccaccagg tctaccaata	5400	
cacgcaaga ag tactatgacg agatcatcaa tgcccttggag gaggatctg cggccaggaa	5460	
gatgcagctg gccttccccc tgacagat tgccgctgca ctggagaaca aggtcactga	5520	
cctctgaccc acaatctcca gtgctgcctt gggacatagg tacctgaggt acctgagagc	5580	
ccctcagggg aggaggccg gtggctgtgg ctgaggcccc caccctcccc tggAACGCGC	5640	
cccaagccgg agtgggtgca gccgaaaccc gcccagcgtc tagactgtag catcttcctc	5700	
tgagcaatac cgccgggac cgcaccagca ccagccccag cccagctcc ctccggccgc	5760	
agaaccagca tcgggtgttc actgtcgagt ctcgagtgat ttgaaaatgt gcctta cgct	5820	
gcccacgtgg gggcagctgg cctccgcctc cggccacgca ccacgccc cctccatgccc	5880	
ctagggtggg cccctgggg atctgaggc ctgtggcccc caggcgaatg tcccgatcc	5940	
tatgtctgtc tgtccaccac gagatgggg gaggagaaaa agcggta cga tgcccttcctg	6000	
acctaaccgg cctccccaag ggtggccgca ctctgggtgg actcacggct gctggccccc	6060	
acgtcaaagg tcaagtgaga cgttagtcaa gtcttacg tggcccccaga catcctgggg	6120	
tcctggctcg tcagacaggg tgccttagag ccccacccag tccggggggg ctgggagcag	6180	
ttccaagacc accccacccc ttttgtaaa tcttgttcat tgtaaatcaa atacagcgtc	6240	
tttttcaactc cg	6252	

<210> 2
<211> 1838
<212> PRT
<213> HOMO SAPIEN

<400> 2
Met Ala Leu Gln Leu Trp Ala Leu Thr Leu Leu Gly Leu Leu Gly Ala
1 5 10 15
Gly Ala Ser Leu Arg Pro Arg Lys Leu Asp Phe Phe Arg Ser Glu Lys
20 25 30
Glu Leu Asn His Leu Ala Val Asp Glu Ala Ser Gly Val Val Tyr Leu
35 40 45
Gly Ala Val Asn Ala Leu Tyr Gln Leu Asp Ala Lys Leu Gln Leu Glu
50 55 60
Gln Gln Val Ala Thr Gly Pro Ala Leu Asp Asn Lys Lys Cys Thr Pro
65 70 75 80
Pro Ile Glu Ala Ser Gln Cys His Glu Ala Glu Met Thr Asp Asn Val
85 90 95
Asn Gln Leu Leu Leu Asp Pro Pro Arg Lys Arg Leu Val Glu Cys
100 105 110
Gly Ser Leu Phe Lys Gly Ile Cys Ala Leu Arg Ala Leu Ser Asn Ile
115 120 125
Ser Leu Arg Leu Phe Tyr Glu Asp Gly Ser Gly Glu Lys Ser Phe Val
130 135 140
Ala Ser Asn Asp Glu Gly Val Ala Thr Val Gly Leu Val Ser Ser Thr
145 150 155 160
Gly Pro Gly Gly Asp Arg Val Leu Phe Val Gly Lys Gly Asn Gly Pro
165 170 175
His Asp Asn Gly Ile Ile Val Ser Thr Arg Leu Leu Asp Arg Thr Asp
180 185 190
Ser Arg Glu Ala Phe Glu Ala Tyr Thr Asp His Ala Thr Tyr Lys Ala

195	200	205													
Gly	Tyr	Leu	Ser	Thr	Asn	Thr	Gln	Gln	Phe	Val	Ala	Ala	Phe	Glu	Asp
210						215				220					
Gly	Pro	Tyr	Val	Phe	Phe	Val	Phe	Asn	Gln	Gln	Asp	Lys	His	Pro	Ala
225						230				235					240
Arg	Asn	Arg	Thr	Leu	Leu	Ala	Arg	Met	Cys	Arg	Glu	Asp	Pro	Asn	Tyr
						245			250						255
Tyr	Ser	Tyr	Leu	Glu	Met	Asp	Leu	Gln	Cys	Arg	Asp	Pro	Asp	Ile	His
						260			265						270
Ala	Ala	Ala	Phe	Gly	Thr	Cys	Leu	Ala	Ala	Ser	Val	Ala	Ala	Pro	Gly
						275			280						285
Ser	Gly	Arg	Val	Leu	Tyr	Ala	Val	Phe	Ser	Arg	Asp	Ser	Arg	Ser	Ser
						290			295						300
Gly	Gly	Pro	Gly	Ala	Gly	Leu	Cys	Leu	Phe	Pro	Leu	Asp	Lys	Val	His
305						310				315					320
Ala	Lys	Met	Glu	Ala	Asn	Arg	Asn	Ala	Cys	Tyr	Thr	Gly	Thr	Arg	Glu
						325			330						335
Ala	Arg	Asp	Ile	Phe	Tyr	Lys	Pro	Phe	His	Gly	Asp	Ile	Gln	Cys	Gly
						340			345						350
Gly	His	Ala	Pro	Gly	Ser	Ser	Lys	Ser	Phe	Pro	Cys	Gly	Ser	Glu	His
						355			360						365
Leu	Pro	Tyr	Pro	Leu	Gly	Ser	Arg	Asp	Gly	Leu	Arg	Gly	Thr	Ala	Val
						370			375						380
Leu	Gln	Arg	Gly	Gly	Leu	Asn	Leu	Thr	Ala	Val	Thr	Val	Ala	Ala	Glu
385						390				395					400
Asn	Asn	His	Thr	Val	Ala	Phe	Leu	Gly	Thr	Ser	Asp	Gly	Arg	Ile	Leu
						405				410					415
Lys	Val	Tyr	Leu	Thr	Pro	Asp	Gly	Thr	Ser	Ser	Glu	Tyr	Asp	Ser	Ile
						420			425						430
Leu	Val	Glu	Ile	Asn	Lys	Arg	Val	Lys	Arg	Asp	Leu	Val	Leu	Ser	Gly
						435			440						445
Asp	Leu	Gly	Ser	Leu	Tyr	Ala	Met	Thr	Gln	Asp	Lys	Val	Phe	Arg	Leu
						450			455						460
Pro	Val	Gln	Glu	Cys	Leu	Ser	Tyr	Pro	Thr	Cys	Thr	Gln	Cys	Arg	Asp
465						470				475					480
Ser	Gln	Asp	Pro	Tyr	Cys	Gly	Trp	Cys	Val	Val	Glu	Gly	Arg	Cys	Thr
						485			490						495
Arg	Lys	Ala	Glu	Cys	Pro	Arg	Ala	Glu	Glu	Ala	Ser	His	Trp	Leu	Trp
						500			505						510
Ser	Arg	Ser	Lys	Ser	Cys	Val	Ala	Val	Thr	Ser	Ala	Gln	Pro	Gln	Asn
						515			520						525
Met	Ser	Arg	Arg	Ala	Gln	Gly	Glu	Val	Gln	Leu	Thr	Val	Ser	Pro	Leu
						530			535						540
Pro	Ala	Leu	Ser	Glu	Glu	Asp	Glu	Leu	Leu	Cys	Leu	Phe	Gly	Glu	Ser
545						550				555					560
Pro	Pro	His	Pro	Ala	Arg	Val	Glu	Gly	Glu	Ala	Val	Ile	Cys	Asn	Ser
						565			570						575
Pro	Ser	Ser	Ile	Pro	Val	Thr	Pro	Gly	Gln	Asp	His	Val	Ala	Val	
						580			585						590
Thr	Ile	Gln	Leu	Leu	Leu	Arg	Arg	Gly	Asn	Ile	Phe	Leu	Thr	Ser	Tyr
						595			600						605
Gln	Tyr	Pro	Phe	Tyr	Asp	Cys	Arg	Gln	Ala	Met	Ser	Leu	Glu	Glu	Asn
						610			615						620
Leu	Pro	Cys	Ile	Ser	Cys	Val	Ser	Asn	Arg	Trp	Thr	Cys	Gln	Trp	Asp
						625			630						640

Leu Arg Tyr His Glu Cys Arg Glu Ala Ser Pro Asn Pro Glu Asp Gly
 645 650 655
 Ile Val Arg Ala His Met Glu Asp Ser Cys Pro Gln Phe Leu Gly Pro
 660 665 670
 Ser Pro Leu Val Ile Pro Met Asn His Glu Thr Asp Val Asn Phe Gln
 675 680 685
 Gly Lys Asn Leu Asp Thr Val Lys Gly Ser Ser Leu His Val Gly Ser
 690 695 700
 Asp Leu Leu Lys Phe Met Glu Pro Val Thr Met Gln Glu Ser Gly Thr
 705 710 715 720
 Phe Ala Phe Arg Thr Pro Lys Leu Ser His Asp Ala Asn Glu Thr Leu
 725 730 735
 Pro Leu His Leu Tyr Val Lys Ser Tyr Gly Lys Asn Ile Asp Ser Lys
 740 745 750
 Leu His Val Thr Leu Tyr Asn Cys Ser Phe Gly Arg Ser Asp Cys Ser
 755 760 765
 Leu Cys Arg Ala Ala Asn Pro Asp Tyr Arg Cys Ala Trp Cys Gly Gly
 770 775 780
 Gln Ser Arg Cys Val Tyr Glu Ala Leu Cys Asn Thr Thr Ser Glu Cys
 785 790 795 800
 Pro Pro Pro Val Ile Thr Arg Ile Gln Pro Glu Thr Gly Pro Leu Gly
 805 810 815
 Gly Gly Ile Arg Ile Thr Ile Leu Gly Ser Asn Leu Gly Val Gln Ala
 820 825 830
 Gly Asp Ile Gln Arg Ile Ser Val Ala Gly Arg Asn Cys Ser Phe Gln
 835 840 845
 Pro Glu Arg Tyr Ser Val Ser Thr Arg Ile Val Cys Val Ile Glu Ala
 850 855 860
 Ala Glu Thr Pro Phe Thr Gly Gly Val Glu Val Asp Val Phe Gly Lys
 865 870 875 880
 Leu Gly Arg Ser Pro Pro Asn Val Gln Phe Thr Phe Gln Gln Pro Lys
 885 890 895
 Pro Leu Ser Val Glu Pro Gln Gln Gly Pro Gln Ala Gly Gly Thr Thr
 900 905 910
 Leu Thr Ile His Gly Thr His Leu Asp Thr Gly Ser Gln Glu Asp Val
 915 920 925
 Arg Val Thr Leu Asn Gly Val Pro Cys Lys Val Thr Lys Phe Gly Ala
 930 935 940
 Gln Leu Gln Cys Val Thr Gly Pro Gln Ala Thr Arg Gly Gln Met Leu
 945 950 955 960
 Leu Glu Val Ser Tyr Gly Gly Ser Pro Val Pro Asn Pro Gly Ile Phe
 965 970 975
 Phe Thr Tyr Arg Glu Asn Pro Val Leu Arg Ala Phe Glu Pro Leu Arg
 980 985 990
 Ser Phe Ala Ser Gly Gly Arg Ser Ile Asn Val Thr Gly Gln Gly Phe
 995 1000 1005
 Ser Leu Ile Gln Arg Phe Ala Met Val Val Ile Ala Glu Pro Leu Gln
 1010 1015 1020
 Ser Trp Gln Pro Pro Arg Glu Ala Glu Ser Leu Gln Pro Met Thr Val
 1025 1030 1035 1040
 Val Gly Thr Asp Tyr Val Phe His Asn Asp Thr Lys Val Val Phe Leu
 1045 1050 1055
 Ser Pro Ala Val Pro Glu Glu Pro Glu Ala Tyr Asn Leu Thr Val Leu
 1060 1065 1070
 Ile Glu Met Asp Gly His Arg Ala Leu Leu Arg Thr Glu Ala Gly Ala

1075	1080	1085
Phe Glu Tyr Val Pro Asp Pro	Thr Phe Glu Asn Phe Thr Gly Gly Val	
1090	1095	1100
Lys Lys Gln Val Asn Lys Leu Ile His Ala Arg Gly Thr Asn Leu Asn		
1105	1110	1115
Lys Ala Met Thr Leu Gln Glu Ala Glu Ala Phe Val Gly Ala Glu Arg		1120
1125	1130	1135
Cys Thr Met Lys Thr Leu Thr Glu Thr Asp Leu Tyr Cys Glu Pro Pro		
1140	1145	1150
Glu Val Gln Pro Pro Pro Lys Arg Arg Gln Lys Arg Asp Thr Thr His		
1155	1160	1165
Asn Leu Pro Glu Phe Ile Val Lys Phe Gly Ser Arg Glu Trp Val Leu		
1170	1175	1180
Gly Arg Val Glu Tyr Asp Thr Arg Val Ser Asp Val Pro Leu Ser Leu		
1185	1190	1195
Ile Leu Pro Leu Val Ile Val Pro Met Val Val Val Ile Ala Val Ser		1200
1205	1210	1215
Val Tyr Cys Tyr Trp Arg Lys Ser Gln Gln Ala Glu Arg Glu Tyr Glu		
1220	1225	1230
Lys Ile Lys Ser Gln Leu Glu Gly Leu Glu Glu Ser Val Arg Asp Arg		
1235	1240	1245
Cys Lys Lys Glu Phe Thr Asp Leu Met Ile Glu Met Glu Asp Gln Thr		
1250	1255	1260
Asn Asp Val His Glu Ala Gly Ile Pro Val Leu Asp Tyr Lys Thr Tyr		
1265	1270	1275
Thr Asp Arg Val Phe Phe Leu Pro Ser Lys Asp Gly Asp Lys Asp Val		
1285	1290	1295
Met Ile Thr Gly Lys Leu Asp Ile Pro Glu Pro Arg Arg Pro Val Val		
1300	1305	1310
Glu Gln Ala Leu Tyr Gln Phe Ser Asn Leu Leu Asn Ser Lys Ser Phe		
1315	1320	1325
Leu Ile Asn Phe Ile His Thr Leu Glu Asn Gln Arg Glu Phe Ser Ala		
1330	1335	1340
Arg Ala Lys Val Tyr Phe Ala Ser Leu Leu Thr Val Ala Leu His Gly		
1345	1350	1355
Lys Leu Glu Tyr Tyr Asp Ile Met His Thr Leu Phe Leu Glu Leu		
1365	1370	1375
Leu Glu Gln Tyr Val Val Ala Lys Asn Pro Lys Leu Met Leu Arg Arg		
1380	1385	1390
Ser Glu Thr Val Val Glu Arg Met Leu Ser Asn Trp Met Ser Ile Cys		
1395	1400	1405
Leu Tyr Gln Tyr Leu Lys Asp Ser Ala Gly Glu Pro Leu Tyr Lys Leu		
1410	1415	1420
Phe Lys Ala Ile Lys His Gln Val Glu Lys Gly Pro Val Asp Ala Val		
1425	1430	1435
Gln Lys Lys Ala Lys Tyr Thr Leu Asn Asp Thr Gly Leu Leu Gly Asp		
1445	1450	1455
Asp Val Glu Tyr Ala Pro Leu Thr Val Ser Val Ile Val Gln Asp Glu		
1460	1465	1470
Gly Val Asp Ala Ile Pro Val Lys Val Leu Asn Cys Asp Thr Ile Ser		
1475	1480	1485
Gln Val Lys Glu Lys Ile Ile Asp Gln Val Tyr Arg Gly Gln Pro Cys		
1490	1495	1500
Ser Cys Trp Pro Arg Pro Asp Ser Val Val Leu Glu Trp Arg Pro Gly		
1505	1510	1515
		1520

Ser Thr Ala Gln Ile Leu Ser Asp Leu Asp Leu Thr Ser Gln Arg Glu
 1525 1530 1535
 Gly Arg Trp Lys Arg Val Asn Thr Leu Met His Tyr Asn Val Arg Asp
 1540 1545 1550
 Gly Ala Thr Leu Ile Leu Ser Lys Val Gly Val Ser Gln Gln Pro Glu
 1555 1560 1565
 Asp Ser Gln Gln Asp Leu Pro Gly Glu Arg His Ala Leu Leu Glu Glu
 1570 1575 1580
 Glu Asn Arg Val Trp His Leu Val Arg Pro Thr Asp Glu Val Asp Glu
 1585 1590 1595 1600
 Gly Lys Ser Lys Arg Gly Ser Val Lys Glu Lys Glu Arg Thr Lys Ala
 1605 1610 1615
 Ile Thr Glu Ile Tyr Leu Thr Arg Leu Leu Ser Val Lys Gly Thr Leu
 1620 1625 1630
 Gln Gln Phe Val Asp Asn Phe Phe Gln Ser Val Leu Ala Pro Gly His
 1635 1640 1645
 Ala Val Pro Pro Ala Val Lys Tyr Phe Asp Phe Leu Asp Glu Gln
 1650 1655 1660
 Ala Glu Lys His Asn Ile Gln Asp Glu Asp Thr Ile His Ile Trp Lys
 1665 1670 1675 1680
 Thr Asn Ser Leu Pro Leu Arg Phe Trp Val Asn Ile Leu Lys Asn Pro
 1685 1690 1695
 His Phe Ile Phe Asp Val His Val His Glu Val Val Asp Ala Ser Leu
 1700 1705 1710
 Ser Val Ile Ala Gln Thr Phe Met Asp Ala Cys Thr Arg Thr Glu His
 1715 1720 1725
 Lys Leu Ser Arg Asp Ser Pro Ser Asn Lys Leu Leu Tyr Ala Lys Glu
 1730 1735 1740
 Ile Ser Thr Tyr Lys Lys Met Val Glu Asp Tyr Tyr Lys Gly Ile Arg
 1745 1750 1755 1760
 Gln Met Val Gln Val Ser Asp Gln Asp Met Asn Thr His Leu Ala Glu
 1765 1770 1775
 Ile Ser Arg Ala His Thr Asp Ser Leu Asn Thr Leu Val Ala Leu His
 1780 1785 1790
 Gln Leu Tyr Gln Tyr Thr Gln Lys Tyr Tyr Asp Glu Ile Ile Asn Ala
 1795 1800 1805
 Leu Glu Glu Asp Pro Ala Ala Gln Lys Met Gln Leu Ala Phe Arg Leu
 1810 1815 1820
 Gln Gln Ile Ala Ala Ala Leu Glu Asn Lys Val Thr Asp Leu
 1825 1830 1835

<210> 3
 <211> 5367
 <212> DNA
 <213> HOMO SAPIEN

<400> 3
 atggctcgct ggcctccctt cggcctctgc ctccctcctgc tgctgctgtc cccaccgcca 60
 ctgcccttga caggggcccac tcgccttcgc gcacctaata ccactctcaa ccacttgcca 120
 ctggcacctg gccgaggcac actctatgtc ggccgcagtga accgcctctt ccagctcagc 180
 cccgagctgc agctcgaggc cgtggctgtc actggccctg taatcgacag ccctgactgc 240
 gtgcccttcc gtgacccagc cgagtccccca caggcccagc tcactgacaa tgccaaaccag 300
 ctgctgctgg tgagcagggc cgcccaggag ctggtgccct gccggcaggt gccggcaggc 360
 gtgtgtgaga cacggcgccct tggggatgtg gccgagggtgc tgtaccaggc tgaggaccct 420
 ggtgacgggc agtttgtggc tgccaataacc ccgggaggtgg caacgggtggg gctggtggtg 480

cccttgccccg	gccgggacact	cctgtttgtg	gccagaggcc	tggcgggcaa	gctgtcgga	540
gggggtgcac	ccctggccat	ccggccagctg	gccgggtctc	agcccttctc	cagcgagggc	600
ctggggccgc	tggtgttggg	cgacttctcc	gactacaaca	acagctacgt	cgggggcctt	660
gccgaacgccc	gctccgccta	cttcgttgc	cgccgcgcg	ggggccgggc	ccaggctgag	720
taccgtctt	acgtggcccg	cgtctgcctg	ggggacacca	acctgtactc	ctacgtggag	780
gtccccctcg	cctgcccagg	ccaggccctc	atccaggccg	ccttccttgc	cccgggcacc	840
ttgcgttaggg	tgtttgcgc	ggggccaaagg	ggcacccagg	cggcgctctg	tgccttcccc	900
atgggtggag	tgggtgcag	cattggagcag	gcccggagac	tctgtacac	ggcgggcggc	960
cggggcccca	gcggcgca	gaaagccacc	gtggagatcg	gcgtcacgtc	gcgctgcgtc	1020
acccctggccc	ttgattcccc	cgagtgcgtac	ccctgtggcg	acagagcacac	ccccagcccc	1080
attgtgtggc	gccagcccc	ggaggtccag	cctctgtga	agctcgggca	gcccggtcagc	1140
gccgtggcag	ctctccaggc	agatgggcac	atgatagcct	tcctggggga	cacccagggc	1200
cagctgtaca	aggctttct	ccacggctcc	cagggccagg	tttaccactc	ccagcaagtg	1260
gggcctccag	gctcagccat	cagggccagac	ctgctgtgg	acagcagtgg	cagtccaccc	1320
tatgtcctga	ctgcccacca	gggtggacgg	ataccgttg	cggctggcc	ccagttccct	1380
gactgtgcca	gctgcccata	ggcccaggac	ccgctgtgt	gctgtgtgt	cctccaggbc	1440
agggttaccc	ggaaggggca	gtgcggggcg	gcaggccagc	tgaaccagtg	gctgtggagt	1500
tatgaggagg	acagccactg	cctgcacatc	cagagcctgc	tgccgggcca	ccaccccccgc	1560
caggagcagg	gccagggtcac	tttgtctgtc	ccccggctgc	ccatcctgga	tgcagatgaa	1620
tacttccatt	gtgcgttcgg	ggactatgac	agcttggctc	atgtggaagg	gccccacgtg	1680
gcctgtgtca	ccccctcccc	agaccaggtg	ccacttaacc	ctccaggcac	agaccacacgtc	1740
actgtgtccc	tggccctgtat	gttcgaggac	gtgactgtgg	ctgcccaccaa	cttctccctt	1800
tatgactgca	gtgcccgtcca	ggccttggag	gcccgtccc	ccgtccttcc	ccagggcctg	1860
cctgcctct	tccactgtcg	gctggagatcg	cctggagaac	ttcggggact	gcccggccacc	1920
ctggaggaga	cagcagggga	ttaaggccctc	atccactgccc	aggcccacca	gccccggagctc	1980
ccagtggcca	tctacgtcac	ccagggtgaa	gcccagagc	tgcacaacac	ccatgtctt	2040
tatggtgagc	ctgagggcag	ccaggcaggg	ggggcaggg	gggtggcaga	caggaggcgc	2100
tcagcacact	gcctgaccc	ccctagtgtat	cctgtacgac	tgcgcacatgg	gccaccccgga	2160
ctgcagccac	tgccaaaggcg	ccaaacaggag	cctgggtctgc	ctgtgaccag	ccctggggccca	2220
ggccccccaa	cccccagcgc	tcggcctggc	tgggctgggt	ggctggccgg	gcaccaggca	2280
ctgcagagtg	gagcgtgggt	gcgggggacc	ccatctgcca	tcatattgcct	gtgcaggtc	2340
gagcccccgt	ccggcccccc	tgagggaggg	ttggccctca	ccatcctggg	ctccaaacctg	2400
ggccgggccc	tcgcccgtgt	gcagtaacgtc	gaccctgtcc	tgctgagcc	gagtccctcgc	2460
tggggcccc	aggcaggggg	cacccagetc	accatccgag	gtcagcaccc	ccagacaggt	2520
ggcaacacca	gtgccttcgt	gggtggccaa	ccctgtcccc	tgggtggcg	actgatccgt	2580
gtcaaggggca	ccggcctaga	cgtgggtcag	cggcccttac	tgtctgtgt	gctggaggct	2640
gacgcagagg	tgcaggcttc	caggggcccg	ccccagacc	cacagccaa	gaggagctgt	2700
ggagcccccgt	ctgcgggacc	ccaggcttgc	atccagctcg	gtggggggct	gtgcagcgc	2760
acagcagagc	ccagctcaact	ccacctgtgg	tcggccctga	atgccccaca	gtgctccacc	2820
gtctgtccg	tcaactgtc	cacccctctc	ctgtgcggg	gcctgtctgt	accagacaga	2880
gcccaacccgc	agcgggtctt	cttcacccta	gacaacgtgc	aagtggactt	cgccagtgcc	2940
agtggggggcc	agggcttctt	gtaccagccc	aaccccccgc	tggcacccct	cagccgcgag	3000
gggcctgtcc	gcccctaccc	cctcaagccca	ggccatgtcc	tggatgtgg	gggcgaggbc	3060
ctcaacccgt	gcatcagcaa	ggaggaggtg	cgcggtcaca	tcggccgcgg	cgagtgcctg	3120
gtgaagacgc	tacacgcac	ccacccgtac	tgcgaggccg	ctgcgcacgc	cccgccagcc	3180
gccaatggct	ccggcctgtcc	acagttcg	gtcgagatgg	gcaatgtca	gctggccctg	3240
ggccctgtc	agtaacgggc	tgaacccccc	ctgtctgcct	ttccctgtgg	ggcccaggca	3300
ggcgtgggca	tgggtgtc	agtgctgatt	gccggcgtgc	tcctccctcac	cctcatgtac	3360
aggcacaaga	gcaaggcaggc	cctgcgggac	taccagaagg	tgcttagtgc	gctggagagc	3420
ctggagacgg	gcgtgggaga	ccagtgcgc	aaggagttca	cagacccat	gacggagatg	3480
accgacacta	gcagcgaccc	ggagggcagc	gggatccct	tcctggacta	ccgcacccctac	3540
gccgagcgc	ccttcttccc	tggccatggc	gttgcggccgc	tgcagcccaa	gcctgägggg	3600
ccaggggagg	acggccactg	tgcactgtg	cgccaggccc	tcacgcagct	ctccaaacctg	3660
ctcaacagca	agctttctt	cctcacggtg	agggccgtgt	ggcgggagtg	cccagtgggc	3720
aaggaggtgg	ggctggggaa	ctactggcct	gagacaaagg	tgggggagga	gacagagacc	3780

atgggtggaga	aactgctcac	caactggctg	tccatctgcc	tgtacgcctt	cctgagggag	3840
gtggctggtg	aaccactgta	catgctttc	cggccatcc	agtaccagg	ggacaaaggc	3900
cccgtggacg	ccgtgacagg	caaggccaaa	cggaccctga	atgatagccg	cttgctgcgg	3960
gaggacgtgg	agttccagcc	cctgacgctg	atggtgcgg	tggggcccg	ggctggcggg	4020
gcccgcaggca	gcagcgagat	gcagcgcgt	ccagccccgg	tgctcgacac	ggacaccatc	4080
acccagggtca	aggagaaggt	gttggaccaa	gtctacaagg	gcacccctt	ctcccagagg	4140
ccctcagtgc	atgcccctaga	ccttggtag	agagccagcc	ctgcccaccc	accccaggga	4200
cccttcccta	cccctccggc	acctggagcc	cctcaactgt	gtcttactat	gaacataccc	4260
acgtgggagg	atggcgagga	ggggggggtg	tgcctctggc	acctggtgaa	agccaccgag	4320
gagccagaag	gggccaaggt	gcccgcggc	agcctgcggg	agcgcgagcc	agcaaggggcc	4380
aaggccattc	cggaaatctt	cctcaccctgt	ctgctgtcca	tgaagggtgg	tgcggcctgg	4440
gtggctgggc	ctgagagggag	gctcagccag	ggaccccgac	cgagccaggg	tgtgggaggg	4500
gcaggggcag	cctcagccgt	ggatggcccc	cacaccctgc	cctccacaca	gcccttatcc	4560
cctgcctcgc	agggcacgct	gcagaagttt	gtggacgaca	ccttccaggc	cattctcagc	4620
gtgaaccggc	ccatccccat	cgcgtcaag	tacctgtttg	accttctgg	tgagctagca	4680
gagaagcacg	gcatcgagga	cccaggacc	ctgcacatct	ggaagaccaa	cagtctgctg	4740
ctgcgttct	gggtgaatgc	cttgaagaac	ccacagctca	tctttgtatgt	acgggtgtcg	4800
gacaatgtgg	acgccccatct	tgcgtcatac	gcccagacct	tcattgactc	ctgtaccacc	4860
tcggagcata	aagtggcccg	ggtgagagca	gtgccagcag	cagcagctgg	caggggcttg	4920
aggagggaaag	gcttatgggg	gaagcctaga	gggctgtgca	cagagctctg	gttgggcaagt	4980
ggcagcatca	tgggggcacc	ttcacctccg	agctcatgcc	tagcgcctcc	cctccctccg	5040
gagcaggatt	ccccagtgaa	caaactgctc	tacgccccgg	agatcccacg	ctacaagcag	5100
atgggtggaga	ggtactatgc	ggacattcgc	cagagctctc	cggcgagcta	ccaggagatg	5160
aactctgtt	tggctgagct	ctccggaaac	tacacttctg	ctccccactg	tctggaggct	5220
ctgcaagaac	tctacaacca	catccacagg	tactatgatc	agattatcag	tgccttggag	5280
gaggaccctg	tggggccagaa	gctgcagctg	gcctgcgc	tgcagcaggt	cggcccccctg	5340
gtggaaaaca	aagtgactga	cctgtga				5367

<210> 4
<211> 1788
<212> PRT
<213> HOMO SAPIEN

<400> 4
Met Ala Arg Trp Pro Pro Phe Gly Leu Cys Leu Leu Leu Leu Leu
1 5 10 15
Ser Pro Pro Pro Leu Pro Leu Thr Gly Ala His Arg Phe Ser Ala Pro
20 25 30
Asn Thr Thr Leu Asn His Leu Ala Leu Ala Pro Gly Arg Gly Thr Leu
35 40 45
Tyr Val Gly Ala Val Asn Arg Leu Phe Gln Leu Ser Pro Glu Leu Gln
50 55 60
Leu Glu Ala Val Ala Val Thr Gly Pro Val Ile Asp Ser Pro Asp Cys
65 70 75 80
Val Pro Phe Arg Asp Pro Ala Glu Cys Pro Gln Ala Gln Leu Thr Asp
85 90 95
Asn Ala Asn Gln Leu Leu Val Ser Ser Arg Ala Gln Glu Leu Val
100 105 110
Ala Cys Gly Gln Val Arg Gln Gly Val Cys Glu Thr Arg Arg Leu Gly
115 120 125
Asp Val Ala Glu Val Leu Tyr Gln Ala Glu Asp Pro Gly Asp Gly Gln
130 135 140
Phe Val Ala Ala Asn Thr Pro Gly Val Ala Thr Val Gly Leu Val Val

145	150	155	160
Pro	Leu	Pro	Gly
Arg	Asp	Leu	Leu
165	170	175	
Lys	Leu	Ser	Ala
Gly	Val	Pro	Pro
180	185	190	
Ser	Gln	Pro	Phe
Ser	Ser	Ser	Glu
Gly	Leu	Gly	Arg
195	200	205	
Phe	Ser	Asp	Tyr
Tyr	Asn	Asn	Ser
210	215	220	
Ser	Ala	Tyr	Phe
Tyr	225	230	235
Arg	240	245	255
Tyr	Arg	Ser	Tyr
Val	Ala	Arg	Val
Cys	Leu	Gly	Asp
260	265	270	
Ser	Tyr	Val	Glu
275	280	285	
Ala	Ala	Phe	Leu
290	295	300	
Pro	Arg	Gly	Thr
Gln	Ala	Ala	Leu
305	310	315	320
Arg	Gly	Pro	Ser
Gly	Ala	Glu	Gly
325	330	335	
Ser	Arg	Cys	Val
340	345	350	
Gly	Asp	Glu	His
355	360	365	
Val	Gln	Pro	Leu
370	375	380	
Leu	Gln	Ala	Asp
385	390	395	400
Gln	Leu	Tyr	Lys
405	410	415	
Ser	Gln	Gln	Val
Gly	Pro	Pro	Gly
420	425	430	
Leu	Asp	Ser	Ser
435	440	445	
Asp	Arg	Ile	Pro
450	455	460	
Cys	Leu	Gln	Ala
465	470	475	480
Arg	Cys	Thr	Arg
485	490	495	
Trp	Leu	Trp	Ser
500	505	510	
Leu	Leu	Pro	Gly
515	520	525	
Ser	Val	Pro	Arg
530	535	540	
Ala	Phe	Gly	Asp
545	550	555	560
Ala	Cys	Val	Thr
565	570	575	
Thr	Asp	His	Val
580	585	590	

Val Ala Ala Thr Asn Phe Ser Phe Tyr Asp Cys Ser Ala Val Gln Ala
 595 600 605
 Leu Glu Ala Ala Ala Pro Val Leu Pro Gln Gly Leu Pro Ala Ser Phe
 610 615 620
 His Cys Trp Leu Glu Leu Pro Gly Glu Leu Arg Gly Leu Pro Ala Thr
 625 630 635 640
 Leu Glu Glu Thr Ala Gly Asp Ser Gly Leu Ile His Cys Gln Ala His
 645 650 655
 Gln Arg Glu Leu Pro Val Pro Ile Tyr Val Thr Gln Gly Glu Ala Gln
 660 665 670
 Arg Leu Asp Asn Thr His Ala Leu Tyr Gly Glu Pro Glu Gly Ser Gln
 675 680 685
 Ala Gly Ala Gly Trp Val Ala Asp Arg Arg Ser Ala His Cys
 690 695 700
 Leu Thr Leu Pro Ser Asp Pro Val Arg Leu Arg His Gly Pro Pro Gly
 705 710 715 720
 Leu Gln Pro Leu Pro Ser Gly Gln Gln Glu Pro Gly Leu Pro Val Thr
 725 730 735
 Ser Pro Ala Pro Gly Pro Gln Thr Pro Ala Ala Arg Pro Gly Trp Ala
 740 745 750
 Gly Trp Leu Ala Gly His Pro Ala Leu Gln Ser Gly Ala Trp Val Arg
 755 760 765
 Gly Thr Pro Ser Ala Ile Ile Cys Leu Leu Gln Val Glu Pro Leu Thr
 770 775 780
 Gly Pro Pro Glu Gly Gly Leu Ala Leu Thr Ile Leu Gly Ser Asn Leu
 785 790 795 800
 Gly Arg Ala Phe Ala Asp Val Gln Tyr Ala Asp Pro Val Leu Leu Ser
 805 810 815
 Leu Ser Pro Arg Trp Gly Pro Gln Ala Gly Gly Thr Gln Leu Thr Ile
 820 825 830
 Arg Gly Gln His Leu Gln Thr Gly Gly Asn Thr Ser Ala Phe Val Gly
 835 840 845
 Gly Gln Pro Cys Pro Met Gly Gly Arg Leu Ile Arg Val Arg Gly Thr
 850 855 860
 Gly Leu Asp Val Val Gln Arg Pro Leu Leu Ser Val Trp Leu Glu Ala
 865 870 875 880
 Asp Ala Glu Val Gln Ala Ser Arg Ala Gln Pro Gln Asp Pro Gln Pro
 885 890 895
 Arg Arg Ser Cys Gly Ala Pro Ala Ala Asp Pro Gln Ala Cys Ile Gln
 900 905 910
 Leu Gly Gly Leu Leu Gln Arg Thr Ala Glu Pro Ser Ser Leu His
 915 920 925
 Leu Trp Ser Ala Leu Asn Ala Pro Gln Cys Ser Thr Val Cys Ser Val
 930 935 940
 Asn Ser Ser Ser Leu Leu Cys Arg Ser Pro Ala Val Pro Asp Arg
 945 950 955 960
 Ala His Pro Gln Arg Val Phe Phe Thr Leu Asp Asn Val Gln Val Asp
 965 970 975
 Phe Ala Ser Ala Ser Gly Gly Gln Gly Phe Leu Tyr Gln Pro Asn Pro
 980 985 990
 Arg Leu Ala Pro Leu Ser Arg Glu Gly Pro Ala Arg Pro Tyr Arg Leu
 995 1000 1005
 Lys Pro Gly His Val Leu Asp Val Glu Gly Glu Gly Leu Asn Leu Gly
 1010 1015 1020
 Ile Ser Lys Glu Glu Val Arg Val His Ile Gly Arg Gly Glu Cys Leu

1025	1030	1035	1040
Val Lys Thr Leu Thr Arg Thr His Leu Tyr Cys Glu Pro Pro Ala His			
1045	1050	1055	
Ala Pro Gln Pro Ala Asn Gly Ser Gly Leu Pro Gln Phe Val Val Gln			
1060	1065	1070	
Met Gly Asn Val Gln Leu Ala Leu Gly Pro Val Gln Tyr Glu Ala Glu			
1075	1080	1085	
Pro Pro Leu Ser Ala Phe Pro Val Glu Ala Gln Ala Gly Val Gly Met			
1090	1095	1100	
Gly Ala Ala Val Leu Ile Ala Ala Val Leu Leu Thr Leu Met Tyr			
1105	1110	1115	1120
Arg His Lys Ser Lys Gln Ala Leu Arg Asp Tyr Gln Lys Val Leu Val			
1125	1130	1135	
Gln Leu Glu Ser Leu Glu Thr Gly Val Gly Asp Gln Cys Arg Lys Glu			
1140	1145	1150	
Phe Thr Asp Leu Met Thr Glu Met Thr Asp Leu Ser Ser Asp Leu Glu			
1155	1160	1165	
Gly Ser Gly Ile Pro Phe Leu Asp Tyr Arg Thr Tyr Ala Glu Arg Ala			
1170	1175	1180	
Phe Phe Pro Gly His Gly Gly Cys Pro Leu Gln Pro Lys Pro Glu Gly			
1185	1190	1195	1200
Pro Gly Glu Asp Gly His Cys Ala Thr Val Arg Gln Gly Leu Thr Gln			
1205	1210	1215	
Leu Ser Asn Leu Leu Asn Ser Lys Leu Phe Leu Leu Thr Val Arg Ala			
1220	1225	1230	
Val Trp Arg Glu Cys Pro Val Gly Lys Glu Val Gly Leu Gly Asn Tyr			
1235	1240	1245	
Trp Pro Glu Thr Lys Val Gly Glu Glu Thr Glu Thr Met Val Glu Lys			
1250	1255	1260	
Leu Leu Thr Asn Trp Leu Ser Ile Cys Leu Tyr Ala Phe Leu Arg Glu			
1265	1270	1275	1280
Val Ala Gly Glu Pro Leu Tyr Met Leu Phe Arg Ala Ile Gln Tyr Gln			
1285	1290	1295	
Val Asp Lys Gly Pro Val Asp Ala Val Thr Gly Lys Ala Lys Arg Thr			
1300	1305	1310	
Leu Asn Asp Ser Arg Leu Leu Arg Glu Asp Val Glu Phe Gln Pro Leu			
1315	1320	1325	
Thr Leu Met Val Leu Val Gly Pro Gly Ala Gly Ala Ala Gly Ser			
1330	1335	1340	
Ser Glu Met Gln Arg Val Pro Ala Arg Val Leu Asp Thr Asp Thr Ile			
1345	1350	1355	1360
Thr Gln Val Lys Glu Lys Val Leu Asp Gln Val Tyr Lys Gly Thr Pro			
1365	1370	1375	
Phe Ser Gln Arg Pro Ser Val His Ala Leu Asp Leu Gly Glu Arg Ala			
1380	1385	1390	
Ser Pro Ala His Pro Pro Gln Gly Pro Phe Pro Thr Pro Pro Ala Pro			
1395	1400	1405	
Gly Ala Pro Gln Leu Cys Leu Thr Met Asn Ile Pro Thr Leu Glu Asp			
1410	1415	1420	
Gly Glu Glu Gly Gly Val Cys Leu Trp His Leu Val Lys Ala Thr Glu			
1425	1430	1435	1440
Glu Pro Glu Gly Ala Lys Val Arg Cys Ser Ser Leu Arg Glu Arg Glu			
1445	1450	1455	
Pro Ala Arg Ala Lys Ala Ile Pro Glu Ile Tyr Leu Thr Arg Leu Leu			
1460	1465	1470	

Ser Met Lys Val Gly Ala Ala Trp Val Ala Gly Pro Glu Arg Arg Leu
 1475 1480 1485
 Ser Gln Gly Pro Arg Pro Ser Gln Gly Val Gly Gly Ala Gly Ala Ala
 1490 1495 1500
 Ser Ala Val Asp Gly Pro His Thr Leu Pro Ser Thr Gln Pro Leu Ser
 1505 1510 1515 1520
 Pro Ala Ser Gln Gly Thr Leu Gln Lys Phe Val Asp Asp Thr Phe Gln
 1525 1530 1535
 Ala Ile Leu Ser Val Asn Arg Pro Ile Pro Ile Ala Val Lys Tyr Leu
 1540 1545 1550
 Phe Asp Leu Leu Asp Glu Leu Ala Glu Lys His Gly Ile Glu Asp Pro
 1555 1560 1565
 Gly Thr Leu His Ile Trp Lys Thr Asn Ser Leu Leu Leu Arg Phe Trp
 1570 1575 1580
 Val Asn Ala Leu Lys Asn Pro Gln Leu Ile Phe Asp Val Arg Val Ser
 1585 1590 1595 1600
 Asp Asn Val Asp Ala Ile Leu Ala Val Ile Ala Gln Thr Phe Ile Asp
 1605 1610 1615
 Ser Cys Thr Thr Ser Glu His Lys Val Gly Arg Val Arg Ala Val Pro
 1620 1625 1630
 Ala Ala Ala Gly Arg Gly Leu Arg Arg Lys Gly Leu Trp Gly Lys
 1635 1640 1645
 Pro Arg Gly Leu Cys Thr Glu Leu Trp Val Gly Ser Gly Ser Ile Met
 1650 1655 1660
 Gly Ala Pro Ser Pro Pro Ser Ser Cys Leu Ala Pro Pro Leu Pro Pro
 1665 1670 1675 1680
 Glu Gln Asp Ser Pro Val Asn Lys Leu Leu Tyr Ala Arg Glu Ile Pro
 1685 1690 1695
 Arg Tyr Lys Gln Met Val Glu Arg Tyr Tyr Ala Asp Ile Arg Gln Ser
 1700 1705 1710
 Ser Pro Ala Ser Tyr Gln Glu Met Asn Ser Ala Leu Ala Glu Leu Ser
 1715 1720 1725
 Gly Asn Tyr Thr Ser Ala Pro His Cys Leu Glu Ala Leu Gln Glu Leu
 1730 1735 1740
 Tyr Asn His Ile His Arg Tyr Tyr Asp Gln Ile Ile Ser Ala Leu Glu
 1745 1750 1755 1760
 Glu Asp Pro Val Gly Gln Lys Leu Gln Leu Ala Cys Arg Leu Gln Gln
 1765 1770 1775
 Val Ala Ala Leu Val Glu Asn Lys Val Thr Asp Leu
 1780 1785

<210> 5
 <211> 5892
 <212> DNA
 <213> HOMO SAPIEN

<400> 5
 ggcacgcccc ggatggctct tcgcgccgcg ggcggcgcac ccttagcgg cccggccgcc 60
 gctgccagcc ccccgccgtt ccagacgcgc cccgggtgcc cggtgccgtt gctgttgctg 120
 ctgtctctgg gggcgccgcg ggccggcgcc ctggagatcc agcgtcggtt cccctcgccc 180
 acgccccacca acaacttcgc cttggacggc gcggcgggga ccgtgtacct ggccggcgtc 240
 aaccgcctct atcagctgtc gggcgccaac ctgagctgg aggccgaggc ggccgtgggc 300
 ccggtgcggc acagccccgt gtgtcacgtc ccgcagctgc cgcaaggctc gtgcgagcac 360
 ccgcggcgcc tcacggacaa ctacaacaag atcctgcagc tggaccccgg ccaggccctg 420
 gtatcgatgt gcgggtccat ctaccaggc ttctgcccagc tgcggcgccg gggtaacatc 480

tccggccgtgg	ccgtgcgctt	cccggcccgcc	gcccggcccg	ccgagcccg	cacggtgttc	540
cccagcatgc	tgaacgtggc	ggccaaccac	ccgaacgcgt	ccaccgtgg	gctagttctg	600
cctccccgccc	cgggcgcggg	gggcagccgc	ctgctcgtgg	gcccacgta	caccggttac	660
ggcagctcct	tcttcccgcg	caaccgcagc	ctggaggacc	accgcttcga	gaacacgccc	720
gagatcgcca	tccgctccct	ggacacgcgc	ggcgacttgg	ccaagcttt	caccttcgac	780
ctcaaccct	ccgacgacaa	catcctcaag	atcaagcagg	gcccggcgg	gcagcacaag	840
ctgggcttcg	tgagcgcctt	cctgcaccccg	tccgaccggc	cgccgggtgc	acagcttac	900
gcgtacctgg	cgctcaaacag	cgaggcgcgc	gcccggcaca	aggagagcga	ggcgcggagc	960
ctgtggcgc	gcatctgcct	gccccacggc	gcccggcggc	acgccaagaa	gctcaccgag	1020
tccractatcc	agttgggctt	gcagtgcgcg	ggcggcgcgg	gccgcggcga	cctctacagc	1080
cgcctggtgt	cggctttccc	agccccggag	cggcttttg	ctgtcttcga	gcggggccag	1140
gggtcccccg	cgggccggcgc	tgctccggcc	gcactctgcg	ccttccgctt	cgccgacgtg	1200
cgagccgcca	tccgagctgc	gcccggcgc	tgcttcgtgg	aaccggcggc	cgacgtggtg	1260
gcgggtcg	acagcgtgtt	gcaggcgcag	ggaccggcct	ggcgcgcga	gctcaacatc	1320
cagctccagc	cagacgcgt	ggactgtgg	gctgctcacc	tgacgcaccc	gtgtccatc	1380
ctcagcccc	tgaaggccac	gcccgtgttc	cgccggccgg	gcctcacctc	cgtggccgtg	1440
gccagcgtca	acaactacac	agcggcttcc	ctgggcacgg	tcaacggag	gcttctcaag	1500
atcaacactg	acgagagcat	gcaggtgggt	agcaggcggg	ttgtgactgt	ggccttatggg	1560
gagcccggtgc	accatgtcat	gcagtttgac	ccagcagact	ccggttaccc	ttacctgtat	1620
acgtccccacc	agatggccag	ggtaaggtc	gccgcctgca	acgtgcactc	cacctgtggg	1680
gactgcgtgg	gtgcggcgg	cgctactgc	ggctgggtgt	ccctggagac	gcccgtgcacc	1740
ttgcagcagg	actgcaccaa	ttccagccag	cagcatttct	ggaccaggc	cagecaggggc	1800
cccagccgc	gtccctccat	gaccgtctcg	ccttccgaga	tcgatgtgc	ccaggagttac	1860
ccaggcatga	tccctcgat	ctcgggcgc	ctgcccagcc	tcagtgccat	ggagatggcc	1920
tgtactatg	ggaacaacat	ccgcactgt	gctcggttc	caggccctgc	ctttggtcac	1980
cagattgcct	actgcacac	cctgcccgg	gaccagttt	cgcccttccc	ccccaaaccag	2040
gaccacgtga	ctgttgagat	gtctgtggg	gtcaatggc	ggaacatgt	caaggccaat	2100
tccaccatct	acgactgcag	ccgcactgca	caagtgtacc	cccacacagc	ctgttaccagc	2160
tgcctgtcg	cacagtggc	ctgtttctgg	tgacggccgc	agactctcg	tgtttccaaac	2220
cagtctcg	gcgaggcctc	accaaacc	acgaggccctc	agactgccc	ccggaccctg	2280
ctccaccccc	tggcacccgt	gcctaccgg	ggctcccgaga	acatccctgg	ggctctggcc	2340
aacactgcct	tttccaggg	tgcaagccctg	gagtgttagt	ttgggctgg	ggagatcttc	2400
gaggctgtgt	gggtgaatga	gtctgttgta	cgctgtgacc	aggtgggtct	gcacacgacc	2460
cggaaagagcc	agggtttccc	gctcagcc	caactaaagg	ggcggccagc	ccgattccctg	2520
gacagccctg	agcccattgc	agtcatgg	tataactgt	ccatgggcag	ccccgactgt	2580
tcccagtgg	tggccgcg	agacccgg	cacctgtgc	tgtggagtga	tggctgcgc	2640
ctggggggc	ctctgcagcc	catggctgg	acctgcccc	cccccggat	ccgcgcgatt	2700
gagcccccgt	gtggccctt	ggacgggtgg	acccctgtga	ccatccgg	aaggAACCTG	2760
ggccggccgc	tca	ggccacccgg	gtgtggatt	gtgggtgtgg	ctgtgagcca	2820
ctgcctgaca	gatacagg	gtcgaggag	atcgtgtgt	tcacaggggc	agccccagga	2880
ccgcctctcg	gtgtgggtac	cgtgaacgc	tctaaggagg	gcaagtcccg	ggaccggcttc	2940
tcctacgtgc	tgccttcgtt	ccactccctg	gagcctacca	ttggccccaa	ggccggggggc	3000
accaggatca	ccatccatgg	gaatgaccc	catgttaggt	ccgagctcca	ggtcctgggt	3060
aacgacacag	acccctgcac	ggagctgtat	cgccacagata	ccagcatcg	ctgcaccatg	3120
cctgggggg	ccctggccgc	tccgggtct	gtgtgtgt	gttgcgg	tcggggctgc	3180
gtgcacggca	acctcacctt	ctgtatcatg	cagaacccgg	tcatcacgg	catcgtcccc	3240
cggccagcc	ctgtcagtgg	ccgcaggacc	atcacagtgg	ctgggtgg	tttccacatg	3300
gtgcagaatg	tgtccatgg	cggtccaccac	attggccgg	agcccacgt	ctgcaagggtt	3360
ctcaactcca	ccctcatcac	ctggccgtcc	cccggggccc	tgagcaacgc	atcagcgcca	3420
gtggacttct	tcatcaatgg	gcgggcctac	gcagacgagg	tggctgtgg	tgaggagcta	3480
ctggaccccg	aggaggcaca	gcggggcagc	aggttccgc	tggactaccc	ccccaaacc	3540
cagtctctta	cggccaaagag	ggagaagtgg	atcaagcacc	accccccgg	gcctctcacc	3600
ctcgttatcc	acaaggagca	ggacagcc	gggtccaga	gtcacgaga	ccgggtcaag	3660
ataggccaag	taagctgcga	catccagatt	gtctctgaca	gaatcatcca	ctgctcggtc	3720
aacgagtccc	tgggcgccggc	cgtggggcag	ctgcccacatca	caatccagg	agggaaacttc	3780

aaccagacca	tcgcccacact	gcagctgggg	ggcagcgaga	cggccatcat	cgtgtccatc	3840
gtcatctgca	gcgtcctgct	gctgctctcc	gtggtgcccc	tgttgcgtct	ctgtaccaag	3900
agccgacgtg	ctgagcgta	ctggcagaag	acgctgctgc	agatggagga	gatggaatct	3960
cagatccgag	aggaaatccg	caaaggcttc	gctgagctgc	agacagacat	gacagatctg	4020
accaaggagc	tgaaccgcag	ccagggcattc	cccttcctgg	agtataagca	cttcgtgacc	4080
cgcaccttct	tccccaaagt	ttcctccctt	tatgaagagc	gttacgtgct	gccctccca	4140
accctcaact	cccaggcag	ctcccaggca	caggaaaccc	acccactgct	gggagagtg	4200
aagatctctg	agagctggcg	gccccacatg	gaagagggaa	ttagcgtgtt	ctccttcacta	4260
ctcaacaaca	agacttctt	catcgtcttt	gtccacgcgc	tgagcagca	gaaggacttt	4320
gcgggtgcgc	acaggtgcag	cctggcctcg	ctgctgacca	tcgcgtgc	cggcaagctg	4380
gagtaactaca	ccagcatcat	gaaggagctg	ctggtgacc	tcattgacgc	ctcgcccgc	4440
aagaacccca	agctcatgt	gccccgcaca	gagtctgtgg	tggagaagat	gctcaccaac	4500
tggatgtcca	tctgcatgt	cagctgtctg	cgggagacgg	tggggagcc	attcttcctg	4560
ctgctgtgtg	ccatcaagca	gcaaatcaac	aagggctcca	tcgacgccc	cacaggcaag	4620
gcccgttaca	cactcaatga	ggagtgccgt	ctgcgggaga	acatcgaggc	caagccccgg	4680
aacctgaacg	tgtccttca	gggctgtggc	atggactcgc	tgacgtgc	gccccatggac	4740
accgacacgc	tgacacagg	caaggagaag	atccctggagg	cttctgc	gaatgtgccc	4800
tactcccagt	ggccgcgtgc	agaggacgtc	gaccttgagt	ggttcgc	cagcacacag	4860
agctacatcc	ttcgggacct	ggacgacacc	tcagtgg	aagacggcc	caagaagctt	4920
aacacgctgg	cccattacaa	gatccctgaa	ggtgcctccc	tggccatgag	tctcatagac	4980
aagaaggaca	acacactggg	ccgagtgaaa	gacttggaca	cagagaagta	tttccatttg	5040
gtgctgccta	cgAACGAGCT	gggggagccc	aagaagtctc	acccggcagag	ccatcgcaag	5100
aagggtctcc	cgggaaatcta	cctgaccgc	ctgctctcca	ccaaagggcac	gttgcagaag	5160
tttctggatg	acctgttcaa	ggcattctg	atgtccctgt	aagacaagcc	cccactggct	5220
gtcaagtaact	ttttcgactt	cttgcggagg	caggctgaga	agagggaaat	ctccgacc	5280
gacaccctac	acatctggaa	gaccaacagc	cttcctctcc	ggttctgggt	gaacatcctg	5340
aagaacccca	agtttgtctt	tgacatcgac	aagacagacc	acatcgacgc	ctgccttca	5400
gtcatcgcc	aggccttcat	cgacgcctgc	tccatctctg	acctgcagct	ggcaaggat	5460
tcgccaacca	acaagctcct	ctacgccaag	gagattcctg	agtaccggaa	gatcg	5520
cgctactaca	agcagatcca	ggacatgacg	ccgctcagcg	agcaagagat	gaatgccc	5580
ctggccgagg	agtcgaggaa	ataccagaat	gagttcaaca	ccaatgtggc	catggcagag	5640
atttttagt	cggccaaagag	gtatcgccg	cagatcatgg	ccgcgcgtgg	ggccaaccc	5700
acggcccgga	ggacacaact	gcagcacaag	tttgagcagg	tggccgtt	gatggaggac	5760
aacatctacg	agtgtacag	tgaggcctga	gacacatgga	gagttggtca	ggctgtct	5820
gggagaaatg	gacgcccact	gggcctcaac	ttgatcttct	accccg	tgtgactcag	5880
actggaaat	ac					5892

<210> 6
<211> 1925
<212> PRT
<213> HOMO SAPIEN

<400> 6
Met Ala Leu Arg Ala Ala Gly Gly Ala Pro Phe Ser Gly Pro Ala Ala
1 5 10 15
Ala Ala Ser Pro Pro Pro Phe Gln Thr Pro Pro Arg Cys Pro Val Pro
20 25 30
Leu Leu Leu Leu Leu Leu Gly Ala Ala Arg Ala Gly Ala Leu Glu
35 40 45
Ile Gln Arg Arg Phe Pro Ser Pro Thr Pro Thr Asn Asn Phe Ala Leu
50 55 60
Asp Gly Ala Ala Gly Thr Val Tyr Leu Ala Ala Val Asn Arg Leu Tyr
65 70 75 80

Gln Leu Ser Gly Ala Asn Leu Ser Leu Glu Ala Glu Ala Ala Val Gly
 85 90 95
 Pro Val Pro Asp Ser Pro Leu Cys His Ala Pro Gln Leu Pro Gln Ala
 100 105 110
 Ser Cys Glu His Pro Arg Arg Leu Thr Asp Asn Tyr Asn Lys Ile Leu
 115 120 125
 Gln Leu Asp Pro Gly Gln Gly Leu Val Val Val Cys Gly Ser Ile Tyr
 130 135 140
 Gln Gly Phe Cys Gln Leu Arg Arg Gly Asn Ile Ser Ala Val Ala
 145 150 155 160
 Val Arg Phe Pro Pro Ala Ala Pro Pro Ala Glu Pro Val Thr Val Phe
 165 170 175
 Pro Ser Met Leu Asn Val Ala Ala Asn His Pro Asn Ala Ser Thr Val
 180 185 190
 Gly Leu Val Leu Pro Pro Ala Ala Gly Ala Gly Gly Ser Arg Leu Leu
 195 200 205
 Val Gly Ala Thr Tyr Thr Gly Ser Ser Phe Phe Pro Arg Asn
 210 215 220
 Arg Ser Leu Glu Asp His Arg Phe Glu Asn Thr Pro Glu Ile Ala Ile
 225 230 235 240
 Arg Ser Leu Asp Thr Arg Gly Asp Leu Ala Lys Leu Phe Thr Phe Asp
 245 250 255
 Leu Asn Pro Ser Asp Asp Asn Ile Leu Lys Ile Lys Gln Gly Ala Lys
 260 265 270
 Glu Gln His Lys Leu Gly Phe Val Ser Ala Phe Leu His Pro Ser Asp
 275 280 285
 Pro Pro Pro Gly Ala Gln Ser Tyr Ala Tyr Leu Ala Leu Asn Ser Glu
 290 295 300
 Ala Arg Ala Gly Asp Lys Glu Ser Gln Ala Arg Ser Leu Leu Ala Arg
 305 310 315 320
 Ile Cys Leu Pro His Gly Ala Gly Gly Asp Ala Lys Lys Leu Thr Glu
 325 330 335
 Ser Tyr Ile Gln Leu Gly Leu Gln Cys Ala Gly Gly Ala Gly Arg Gly
 340 345 350
 Asp Leu Tyr Ser Arg Leu Val Ser Val Phe Pro Ala Arg Glu Arg Leu
 355 360 365
 Phe Ala Val Phe Glu Arg Pro Gln Gly Ser Pro Ala Ala Arg Ala Ala
 370 375 380
 Pro Ala Ala Leu Cys Ala Phe Arg Phe Ala Asp Val Arg Ala Ala Ile
 385 390 395 400
 Arg Ala Ala Arg Thr Ala Cys Phe Val Glu Pro Ala Pro Asp Val Val
 405 410 415
 Ala Val Leu Asp Ser Val Val Gln Gly Thr Gly Pro Ala Cys Glu Arg
 420 425 430
 Lys Leu Asn Ile Gln Leu Gln Pro Glu Gln Leu Asp Cys Gly Ala Ala
 435 440 445
 His Leu Gln His Pro Leu Ser Ile Leu Gln Pro Leu Lys Ala Thr Pro
 450 455 460
 Val Phe Arg Ala Pro Gly Leu Thr Ser Val Ala Val Ala Ser Val Asn
 465 470 475 480
 Asn Tyr Thr Ala Val Phe Leu Gly Thr Val Asn Gly Arg Leu Leu Lys
 485 490 495
 Ile Asn Leu Asn Glu Ser Met Gln Val Val Ser Arg Arg Val Val Thr
 500 505 510
 Val Ala Tyr Gly Glu Pro Val His His Val Met Gln Phe Asp Pro Ala

	515	520	525
Asp Ser Gly Tyr Leu Tyr	Leu Met Thr Ser His Gln	Met Ala Arg Val	
530	535	540	
Lys Val Ala Ala Cys Asn Val His Ser Thr Cys Gly Asp Cys Val Gly			
545	550	555	560
Ala Ala Asp Ala Tyr Cys Gly Trp Cys Ala Leu Glu Thr Arg Cys Thr			
565	570	575	
Leu Gln Gln Asp Cys Thr Asn Ser Ser Gln Gln His Phe Trp Thr Ser			
580	585	590	
Ala Ser Glu Gly Pro Ser Arg Cys Pro Ala Met Thr Val Leu Pro Ser			
595	600	605	
Glu Ile Asp Val Arg Gln Glu Tyr Pro Gly Met Ile Leu Gln Ile Ser			
610	615	620	
Gly Ser Leu Pro Ser Leu Ser Gly Met Glu Met Ala Cys Asp Tyr Gly			
625	630	635	640
Asn Asn Ile Arg Thr Val Ala Arg Val Pro Gly Pro Ala Phe Gly His			
645	650	655	
Gln Ile Ala Tyr Cys Asn Leu Leu Pro Arg Asp Gln Phe Pro Pro Phe			
660	665	670	
Pro Pro Asn Gln Asp His Val Thr Val Glu Met Ser Val Arg Val Asn			
675	680	685	
Gly Arg Asn Ile Val Lys Ala Asn Phe Thr Ile Tyr Asp Cys Ser Arg			
690	695	700	
Thr Ala Gln Val Tyr Pro His Thr Ala Cys Thr Ser Cys Leu Ser Ala			
705	710	715	720
Gln Trp Pro Cys Phe Trp Cys Ser Gln Gln His Ser Cys Val Ser Asn			
725	730	735	
Gln Ser Arg Cys Glu Ala Ser Pro Asn Pro Thr Ser Pro Gln Asp Cys			
740	745	750	
Pro Arg Thr Leu Leu Ser Pro Leu Ala Pro Val Pro Thr Gly Gly Ser			
755	760	765	
Gln Asn Ile Leu Val Pro Leu Ala Asn Thr Ala Phe Phe Gln Gly Ala			
770	775	780	
Ala Leu Glu Cys Ser Phe Gly Leu Glu Glu Ile Phe Glu Ala Val Trp			
785	790	795	800
Val Asn Glu Ser Val Val Arg Cys Asp Gln Val Val Leu His Thr Thr			
805	810	815	
Arg Lys Ser Gln Val Phe Pro Leu Ser Leu Gln Leu Lys Gly Arg Pro			
820	825	830	
Ala Arg Phe Leu Asp Ser Pro Glu Pro Met Thr Val Met Val Tyr Asn			
835	840	845	
Cys Ala Met Gly Ser Pro Asp Cys Ser Gln Cys Leu Gly Arg Glu Asp			
850	855	860	
Leu Gly His Leu Cys Val Trp Ser Asp Gly Cys Arg Leu Arg Gly Pro			
865	870	875	880
Leu Gln Pro Met Ala Gly Thr Cys Pro Ala Pro Glu Ile Arg Ala Ile			
885	890	895	
Glu Pro Leu Ser Gly Pro Leu Asp Gly Gly Thr Leu Leu Thr Ile Arg			
900	905	910	
Gly Arg Asn Leu Gly Arg Arg Leu Ser Asp Val Ala His Gly Val Trp			
915	920	925	
Ile Gly Gly Val Ala Cys Glu Pro Leu Pro Asp Arg Tyr Thr Val Ser			
930	935	940	
Glu Glu Ile Val Cys Val Thr Gly Pro Ala Pro Gly Pro Leu Ser Gly			
945	950	955	960

Val Val Thr Val Asn Ala Ser Lys Glu Gly Lys Ser Arg Asp Arg Phe
 965 970 975
 Ser Tyr Val Leu Pro Leu Val His Ser Leu Glu Pro Thr Met Gly Pro
 980 985 990
 Lys Ala Gly Gly Thr Arg Ile Thr Ile His Gly Asn Asp Leu His Val
 995 1000 1005
 Gly Ser Glu Leu Gln Val Leu Val Asn Asp Thr Asp Pro Cys Thr Glu
 1010 1015 1020
 Leu Met Arg Thr Asp Thr Ser Ile Ala Cys Thr Met Pro Glu Gly Ala
 1025 1030 1035 1040
 Leu Pro Ala Pro Val Pro Val Cys Val Arg Phe Glu Arg Arg Gly Cys
 1045 1050 1055
 Val His Gly Asn Leu Thr Phe Trp Tyr Met Gln Asn Pro Val Ile Thr
 1060 1065 1070
 Ala Ile Ser Pro Arg Arg Ser Pro Val Ser Gly Gly Arg Thr Ile Thr
 1075 1080 1085
 Val Ala Gly Glu Arg Phe His Met Val Gln Asn Val Ser Met Ala Val
 1090 1095 1100
 His His Ile Gly Arg Glu Pro Thr Leu Cys Lys Val Leu Asn Ser Thr
 1105 1110 1115 1120
 Leu Ile Thr Cys Pro Ser Pro Gly Ala Leu Ser Asn Ala Ser Ala Pro
 1125 1130 1135
 Val Asp Phe Phe Ile Asn Gly Arg Ala Tyr Ala Asp Glu Val Ala Val
 1140 1145 1150
 Ala Glu Glu Leu Leu Asp Pro Glu Glu Ala Gln Arg Gly Ser Arg Phe
 1155 1160 1165
 Arg Leu Asp Tyr Leu Pro Asn Pro Gln Phe Ser Thr Ala Lys Arg Glu
 1170 1175 1180
 Lys Trp Ile Lys His His Pro Gly Glu Pro Leu Thr Leu Val Ile His
 1185 1190 1195 1200
 Lys Glu Gln Asp Ser Leu Gly Leu Gln Ser His Glu Tyr Arg Val Lys
 1205 1210 1215
 Ile Gly Gln Val Ser Cys Asp Ile Gln Ile Val Ser Asp Arg Ile Ile
 1220 1225 1230
 His Cys Ser Val Asn Glu Ser Leu Gly Ala Ala Val Gly Gln Leu Pro
 1235 1240 1245
 Ile Thr Ile Gln Val Gly Asn Phe Asn Gln Thr Ile Ala Thr Leu Gln
 1250 1255 1260
 Leu Gly Gly Ser Glu Thr Ala Ile Ile Val Ser Ile Val Ile Cys Ser
 1265 1270 1275 1280
 Val Leu Leu Leu Ser Val Val Ala Leu Phe Val Phe Cys Thr Lys
 1285 1290 1295
 Ser Arg Arg Ala Glu Arg Tyr Trp Gln Lys Thr Leu Leu Gln Met Glu
 1300 1305 1310
 Glu Met Glu Ser Gln Ile Arg Glu Ile Arg Lys Gly Phe Ala Glu
 1315 1320 1325
 Leu Gln Thr Asp Met Thr Asp Leu Thr Lys Glu Leu Asn Arg Ser Gln
 1330 1335 1340
 Gly Ile Pro Phe Leu Glu Tyr Lys His Phe Val Thr Arg Thr Phe Phe
 1345 1350 1355 1360
 Pro Lys Cys Ser Ser Leu Tyr Glu Glu Arg Tyr Val Leu Pro Ser Gln
 1365 1370 1375
 Thr Leu Asn Ser Gln Gly Ser Ser Gln Ala Gln Glu Thr His Pro Leu
 1380 1385 1390
 Leu Gly Glu Trp Lys Ile Pro Glu Ser Cys Arg Pro Asn Met Glu Glu

1395	1400	1405
Gly Ile Ser Val Phe Ser Ser	Leu Leu Asn Asn Lys His Phe Leu Ile	
1410 1415	1420	
Val Phe Val His Ala Leu Glu Gln Gln Lys Asp Phe Ala Val Arg Asp		
1425 1430	1435	1440
Arg Cys Ser Leu Ala Ser Leu Leu Thr Ile Ala Leu His Gly Lys Leu		
1445 1450	1455	1455
Glu Tyr Tyr Thr Ser Ile Met Lys Glu Leu Leu Val Asp Leu Ile Asp		
1460 1465	1470	
Ala Ser Ala Ala Lys Asn Pro Lys Leu Met Leu Arg Arg Thr Glu Ser		
1475 1480	1485	
Val Val Glu Lys Met Leu Thr Asn Trp Met Ser Ile Cys Met Tyr Ser		
1490 1495	1500	
Cys Leu Arg Glu Thr Val Gly Glu Pro Phe Phe Leu Leu Leu Cys Ala		
1505 1510	1515	1520
Ile Lys Gln Gln Ile Asn Lys Gly Ser Ile Asp Ala Ile Thr Gly Lys		
1525 1530	1535	
Ala Arg Tyr Thr Leu Asn Glu Glu Trp Leu Leu Arg Glu Asn Ile Glu		
1540 1545	1550	
Ala Lys Pro Arg Asn Leu Asn Val Ser Phe Gln Gly Cys Gly Met Asp		
1555 1560	1565	
Ser Leu Ser Val Arg Ala Met Asp Thr Asp Thr Leu Thr Gln Val Lys		
1570 1575	1580	
Glu Lys Ile Leu Glu Ala Phe Cys Lys Asn Val Pro Tyr Ser Gln Trp		
1585 1590	1595	1600
Pro Arg Ala Glu Asp Val Asp Leu Glu Trp Phe Ala Ser Ser Thr Gln		
1605 1610	1615	
Ser Tyr Ile Leu Arg Asp Leu Asp Asp Thr Ser Val Val Glu Asp Gly		
1620 1625	1630	
Arg Lys Lys Leu Asn Thr Leu Ala His Tyr Lys Ile Pro Glu Gly Ala		
1635 1640	1645	
Ser Leu Ala Met Ser Leu Ile Asp Lys Lys Asp Asn Thr Leu Gly Arg		
1650 1655	1660	
Val Lys Asp Leu Asp Thr Glu Lys Tyr Phe His Leu Val Leu Pro Thr		
1665 1670	1675	1680
Asp Glu Leu Ala Glu Pro Lys Lys Ser His Arg Gln Ser His Arg Lys		
1685 1690	1695	
Lys Val Leu Pro Glu Ile Tyr Leu Thr Arg Leu Leu Ser Thr Lys Gly		
1700 1705	1710	
Thr Leu Gln Lys Phe Leu Asp Asp Leu Phe Lys Ala Ile Leu Ser Ile		
1715 1720	1725	
Arg Glu Asp Lys Pro Pro Leu Ala Val Lys Tyr Phe Phe Asp Phe Leu		
1730 1735	1740	
Glu Glu Gln Ala Glu Lys Arg Gly Ile Ser Asp Pro Asp Thr Leu His		
1745 1750	1755	1760
Ile Trp Lys Thr Asn Ser Leu Pro Leu Arg Phe Trp Val Asn Ile Leu		
1765 1770	1775	
Lys Asn Pro Gln Phe Val Phe Asp Ile Asp Lys Thr Asp His Ile Asp		
1780 1785	1790	
Ala Cys Leu Ser Val Ile Ala Gln Ala Phe Ile Asp Ala Cys Ser Ile		
1795 1800	1805	
Ser Asp Leu Gln Leu Gly Lys Asp Ser Pro Thr Asn Lys Leu Leu Tyr		
1810 1815	1820	
Ala Lys Glu Ile Pro Glu Tyr Arg Lys Ile Val Gln Arg Tyr Tyr Lys		
1825 1830	1835	1840

Gln Ile Gln Asp Met Thr Pro Leu Ser Glu Gln Glu Met Asn Ala His
 1845 1850 1855
 Leu Ala Glu Glu Ser Arg Lys Tyr Gln Asn Glu Phe Asn Thr Asn Val
 1860 1865 1870
 Ala Met Ala Glu Ile Phe Arg Ser Pro Lys Arg Tyr Arg Pro Gln Ile
 1875 1880 1885
 Met Ala Ala Leu Glu Ala Asn Pro Thr Ala Arg Arg Thr Gln Leu Gln
 1890 1895 1900
 His Lys Phe Glu Gln Val Val Ala Leu Met Glu Asp Asn Ile Tyr Glu
 1905 1910 1915 1920
 Cys Tyr Ser Glu Ala
 1925

<210> 7

<211> 601

<212> DNA

<213> HOMO SAPIEN

<400> 7

caccagagtc	cctgtggagt	cctgtggtca	gtatcagagc	tgcggcgagt	gccttggctc	60
aggcgacccc	caactgtggct	ggtgtgtgct	gcacaacact	tgcaccccgga	aggagcggtg	120
tgagcggtcc	aaggagcccc	gcagggttgc	ctcggagatg	aagcagtgtg	tccggctgac	180
ggtcacatccc	aacaataatct	ccgtctctca	gtacaacgag	ctgctggtcc	tggagacgta	240
caatgtcccc	gagctgtcag	ctggcgtcaa	ctgcacctt	gaggacctgt	cagagatgga	300
tgggctggtc	gtgggcaatc	agatccagtg	ctactccct	gcagccaagg	aggtgccccg	360
gatcatcaca	gagaatgggg	accaccaatgt	cgtacagctt	cagctaaat	caaaggagac	420
cggcatgacc	ttcggccagca	ccagctttgt	cttctacaat	tgagcgtcc	acaattcgtg	480
cctgtcctgc	gtggagagtc	cataaccgctg	ccactgtgt	aaataccggc	atgtctgcac	540
ccatgacccc	aagacctgtc	ccttccagga	aggccgagtg	aagctgcccc	agtaggtcc	600
c						601

<210> 8

<211> 199

<212> PRT

<213> HOMO SAPIEN

<400> 8

Thr Arg Val Pro Val Glu Ser Cys Gly Gln Tyr Gln Ser Cys Gly Glu						
1	5	10	15			
Cys Leu Gly Ser Gly Asp Pro His Cys Gly Trp Cys Val Leu His Asn						
20	25	30				
Thr Cys Thr Arg Lys Glu Arg Cys Glu Arg Ser Lys Glu Pro Arg Arg						
35	40	45				
Phe Ala Ser Glu Met Lys Gln Cys Val Arg Leu Thr Val His Pro Asn						
50	55	60				
Asn Ile Ser Val Ser Gln Tyr Asn Ala Leu Leu Val Leu Glu Thr Tyr						
65	70	75	80			
Asn Val Pro Glu Leu Ser Ala Gly Val Asn Cys Thr Phe Glu Asp Leu						
85	90	95				
Ser Glu Met Asp Gly Leu Val Val Gly Asn Gln Ile Gln Cys Tyr Ser						
100	105	110				
Pro Ala Ala Lys Glu Val Pro Arg Ile Ile Thr Glu Asn Gly Asp His						
115	120	125				
His Val Val Gln Leu Gln Leu Lys Ser Lys Glu Thr Gly Met Thr Phe						
130	135	140				

Ala Ser Thr Ser Phe Val Phe Tyr Asn Cys Ser Val His Asn Ser Cys
 145 150 155 160
 Leu Ser Cys Val Glu Ser Pro Tyr Arg Cys His Trp Cys Lys Tyr Arg
 165 170 175
 His Val Cys Thr Asp Pro Lys Thr Cys Ser Phe Gln Glu Gly Arg Val
 180 185 190
 Lys Leu Pro Glu Val Gly Pro
 195

<210> 9

<211> 6408

<212> DNA

<213> HOMO SAPIEN

<400> 9

atgcctgctc	tggcccccagc	tcttctccag	gctctctggg	ccgggtgggt	cctcaccctc	60
cagcccccttc	caccaactgc	attcaactccc	aatggcacgt	atctgcagca	cctggcaagg	120
gaccggcacct	caggcacccct	ctacactgggg	gctaccaaact	tccctgttcca	gctgagccct	180
gggctgcagc	tggaggccac	agtgtccacc	ggccctgtgc	tagacagcag	ggactgcctg	240
ccacactgtga	tgcctgtatga	gtggcccccag	gcccaagcccta	ccaacaaccc	gaatcagctg	300
ctcctgggtga	gcccaggggc	ctctgtggta	tgcgggagcg	tgcaccaggg	ggtctgtgaa	360
cagccggcc	tggggcagct	cgacgacgtc	ctgctgcggc	cagagcggcc	tggggacaca	420
caatatgtgg	ctgccaatgt	tcctgcggtc	agcacgggtgg	ggctgttagc	ccagggcttg	480
gcagggggc	ccctcctgtt	tgtggggcga	ggatacaca	gcaggggtgt	gggggggtggc	540
attccaccca	tcacaacccg	ggccctgtgg	ccgcccggacc	cccaagctgc	tttctccstat	600
gaggagacag	ccaagctggc	agtggggccgc	ctctccggagt	acagccacca	ttcgtgagtt	660
gccttgcac	gtggggccag	cgcctacttc	ctgttcttc	ggcgggaccc	gcaggctcag	720
tctagagctt	ttcgtgccta	tgtatctcga	gtgtgtctcc	gggaccagca	ctactactcc	780
tatgtggagt	tgcctctggc	ctgcgaaggt	ggccgcatacg	ggctgatcca	ggctgcagct	840
gtggccacgt	ccaggggaggt	ggcgcatggg	gaggtctct	ttgcagcttt	ctcctcggct	900
gcacccccc	ctgtggggcg	gccccatctcg	gcccgtctg	gggcatctgg	agcctctgcc	960
ctctgtgcct	tcccccttgg	tgaggtggac	cggcttgcata	atcgcacgcg	agatgcctgc	1020
tacacccggg	agggtcggtc	tgaggatggg	accgagggtgg	cctacatcg	gtatgatgtc	1080
aattctgact	gtgcacagct	gccagtggac	accctggatg	cttacccctg	tggctcagac	1140
cacacgcccc	gccccatggc	cagccgggtc	ccgctggaa	ccacaccaat	tctggagtgg	1200
ccagggattc	agctaacagc	tgtggcagtc	accatggaa	atggacacac	catcgctttc	1260
ctgggtgata	gtcaaggggca	gctgcacagg	gtctacttgg	gcccaggggag	cgatggccac	1320
ccatactcca	cacagagcat	ccagcgggg	tctgcagtg	gcagagaccc	cacctttgat	1380
gggacctttg	agcacctgt	tgtcatgacc	cagagcacac	ttctgaagggt	tcctgtggct	1440
tccttgtc	agcacctgg	ctgtgcata	tgccttgc	acagggaccc	atactgtggg	1500
tggtgtcgtc	tccttggcag	gtcagtcgc	cgttctgagt	gtcggagggg	ccagggccca	1560
gagcagtggc	tatggagctt	ccagcctgag	ctgggcgtc	tgcaagtggc	agccatgagt	1620
cctgccaaca	tcagccgaga	ggagacgagg	gaggtttcc	tatcagtg	ccagaccccca	1680
ccccctgtgc	cagggggatc	atattctgc	cactttgggg	aacatcagag	tcctgccctg	1740
ctgactgggt	ctgggtgtat	gtggccctcc	ccagaccata	gtgaggcccc	agtgtgtccg	1800
agaggagccg	actacgtatc	cgtgagcgtg	gagctcagat	ttggcgctgt	tgtgatcgcc	1860
aaaacttccc	tctctttcta	tgactgtgt	gcccgtca	aactccgccc	atctgcgcag	1920
tgccaggcc	gtgtgagcag	ccgcgtgggg	tgtaactgg	gtgtctggca	gcacctgtgc	1980
accacacaagg	cctcgtgt	tgctggggcc	atggttgaa	gcatcagag	cccgcttgc	2040
tccccagacc	cctctgcaag	agggtggaccc	agccccctcc	cacccacagc	ccccaaagcc	2100
ctggccaccc	ctgctctgt	cacccttccc	gtggagccctg	gggcctccctc	cacagccaca	2160
gcttggaca	tctcacctgg	ggctagtcct	tccctgtca	gccccctgggg	gccatgggca	2220
ggttctggct	ccatatcttc	ccctggctcc	acagggtcgc	ctctccatgt	ggagccctcc	2280
cctcccaagcc	ccccaaatgg	acctggaaacc	gtgtccctg	cccccaactga	ttcagacacc	2340
tcagccacac	ctgaggaccc	cttggccctcc	ccgctgtc	cgtcagaggt	agcagcagtg	2400

ccccctgcag	accctggccc	cgaggctctt	catcccacag	tgccccctgga	cctgccccct	2460
gccactgttc	ctgccaccac	tttcccaggg	gccatgggct	ccgtgaagcc	cgcctggac	2520
tggctcacga	gagaaggcgg	cgagctgccc	gaggcggacg	agtggacggg	gggtgacgca	2580
ccgccttct	ccacttccac	cctcctctca	ggtgatggag	actcagcaga	gcttgagggc	2640
cctcccgccc	ccctcatcct	cccgctccagc	ctcgactacc	agtatgacac	ccccgggctc	2700
tgggagctgg	aagaggcgcac	cttggggca	agctcctgccc	cctgtgtgga	gagcgtttag	2760
ggctccacgt	tgtatgcccgt	ccatgtggag	cgggaaatcc	ggctgctagg	caggaacctg	2820
caccctttcc	aggatggccc	aggagacaat	gagtgtgtga	tggagctgga	gggcctcgag	2880
gtgggttgt	aggccccgggt	cgagtgtgag	ccacccctccag	ataccccatgt	ccatgtcacc	2940
tgcacgcgc	accagctcag	ctatgaggct	ctgcagccgg	agtcctgtgt	ggggctgttt	3000
ctgcgtcggg	ccggccgtct	cgctgtggac	agtgctgagg	ggctgcatgt	gttactgtat	3060
gactgttccg	tgggacatgg	agactgcagc	cgctgcca	ctgccatgcc	ccagtatggc	3120
tgtgtgttgt	gtgaggggga	gcgtccacgt	tgtgtgaccc	gggaggcctg	tggtagggct	3180
gaggctgtgg	ccacccactg	cccagcgc	ctcatccact	cggtggagcc	actgactggg	3240
cctgttagacg	gaggcaccgg	tgtcaccatc	aggggctcca	acctgggcca	gcatgtgcag	3300
gatgtgtcgg	gcattgtcac	ggtggctgg	gtgcccctgt	ctgtggatgc	ccagggatgc	3360
gaggctctcg	gcagcctcg	gtgcattacc	ggggccatgt	gggaggaggt	ggccggcgcc	3420
acagccgtgg	agggtccggg	aaggaggatgt	ggtgcattcg	aacacgactt	tgcctaccag	3480
gatccgaagg	tccattccat	cttcccgccc	cgccggccca	gagctggggg	caccgtctc	3540
accctgaatg	gctccaagct	cctgactggg	cggtggagg	acatccgagt	gttggttgga	3600
gaccagcctt	gtcacttgct	gcccggagcag	cagtcaaaac	aactgcgggt	ttagaccaggc	3660
ccacgccccca	cgccctccac	gctccctgt	gctgtgtgt	ttggggccac	ggagcggagg	3720
cttcaacgcgc	gacagttcaa	gtatacctt	gaccccaaca	tcacctctgc	tggcccccacc	3780
aagagcttcc	tcagtggagg	acgtgagata	tgcgtccgt	gccaagatgt	ggacgtggta	3840
cagacgccaa	gaatccgggt	gaccgtgg	tgcagaatgc	tgcaagccaa	ccagggctt	3900
ggacggaggc	gtcgcgtgt	ccccggagcag	gcattgtccc	tttgcaccc	ctgcagtagc	3960
cagaatattg	aggagccgt	ccatgtcaac	tcctccac	tcatcacgt	ccgcacacct	4020
gccctcccg	gcctgcctga	ggacccctgg	gtccgggtgg	aatttac	tgacaacctg	4080
gtctttgact	ttgcaacact	gaacccca	cttttctct	atgaggccg	ccccaccctg	4140
cagccactca	accctgagga	ccccaccatg	ccattccggc	acaagctgg	gagtgtttc	4200
tccgtggagg	gggagaac	ggacccctg	atgtccaa	aggaggtgt	ggctatgata	4260
ggggatggcc	cctgttgtt	gaagacgt	acgcggcacc	acctgtact	cgagcccccc	4320
gtggagcgc	ccctgcccac	gcacccatgc	ctccggagg	cacgtgact	tttgcctgag	4380
ttcacgggtc	agatggggaa	cttgcgttc	tccctgggt	acgtgcagta	tgacggcgag	4440
agccctgggg	cttttccgt	ggcagccca	gtgggcttgg	gggtggggc	ctctcttctg	4500
gctctgggt	tcatcatcat	tgcctcat	tacaggagga	agagcaagca	ggccctgagg	4560
gactataaga	aggttcagat	ccagctggag	aatctggaga	gcagtgtgc	ggaccgctgc	4620
aagaagaaat	tcacagac	catgactgag	atgaccgatc	tcaccagt	cctctgggc	4680
agcggcattc	ccttcctcg	ctacaagggt	tatgcggaga	ggatcttct	ccctgggcac	4740
cgcgagtcgc	ccttgcac	ggacccctgg	gtgcgttgc	gcagacggc	cactgttagag	4800
caaggcgtgg	ggcagcttc	taacctgtc	aacacgaa	tcttcctc	caagttcatc	4860
cacacgctgg	agacccac	cacccctca	gctcgggacc	gtcctacgt	ggcatctctg	4920
ctcaccgtgg	cactgcatgg	gaagctttag	tatttact	acatccctcc	cactgtctc	4980
agtgacctgg	ttgcccagta	tgtggcca	aacccca	tgtatgcgt	caggacagag	5040
actgtgggt	agaagctgt	caccaactgg	atgtccatct	gtctgtata	ttcgtgtt	5100
gactccctgt	gggagccct	gtacatgtc	tttcgagg	ttaagcacca	agtggataag	5160
ggcccgatgg	acagtgtac	aggcaagg	aaatacac	tgaacgacaa	ccgcctgtc	5220
agagaggatg	tggagtacc	tccctgacc	ttgaatgcac	tattggctgt	ggggcttgg	5280
gcaggagagg	cccaggcgt	gcccgtaa	gtcctagact	gtgacccat	ccccaggca	5340
aaggagaaga	tgctggacca	gcttataaa	ggagtgcctc	tcacccac	gccagaccct	5400
cgcaccctt	atgttgatgt	gcccgttgg	gtggccggc	acctcattt	ttctgacgag	5460
gtgtcactt	ctgaggatcca	gggtctgtt	aggcgcctga	acacactgca	gcattacaag	5520
gtcccagatg	gagcaactgt	ggccctcg	ccctgcctca	ccaagcatgt	gttccggaa	5580
aaccaggatt	atgtccctgg	agagcggacc	ccaatgcgtt	aggatgtaga	tgagggggggc	5640
atccggccct	ggcacctgtt	gaagccaa	gtgagccgg	agccgccc	gcctcgagg	5700

ggcagccttc	ggggcgaaaa	gcgtgagcgc	gccaaaggcca	tccctgagat	ctacctgacc	5760
cgccctgtgt	ccatgaagggg	caccctgcag	aagtcgtgg	atgacactgtt	ccaggtgatt	5820
ctcagcacca	gccgccccgt	gccgctcgct	gtgaagtact	tctttgacct	gctggatgag	5880
caggcccagc	agcatggcat	ctccgaccag	gacaccatcc	acatctggaa	gaccaacagc	5940
ttgcctctga	ggttctggat	caatataata	aaaaacccgc	agtttgtgtt	cgacgtgcaa	6000
acatctgata	acatggatgc	ggtgcctctt	gtcattgcac	agaccttcat	ggacgcctgc	6060
accctggccg	accacaagct	ggggccggac	tccccgatca	acaaaacttct	gtatgcacgg	6120
gacattcccc	ggtacaagcg	gatgggtggaa	aggtactatg	cagacatcaag	acagactgtc	6180
ccagccagcg	accaagagat	gaactctgtc	ctggctgaac	tgtcctggaa	ctactccgg	6240
gacctcgggg	cgcgagtggc	cctgcatgaa	ctctacaagt	acatcaacaa	gtactatgac	6300
cagatcatca	ctgcccctgga	ggaggatggc	acggccaga	agatgcagct	ggctatcgg	6360
ctccagcaga	ttgcagctgc	tgtggaaaac	aaggtcacag	atctatag		6408

<210> 10
<211> 2135
<212> PRT
<213> HOMO SAPIEN

<400> 10
Met Pro Ala Leu Gly Pro Ala Leu Leu Gln Ala Leu Trp Ala Gly Trp
1 5 10 15
Val Leu Thr Leu Gln Pro Leu Pro Pro Thr Ala Phe Thr Pro Asn Gly
20 25 30
Thr Tyr Leu Gln His Leu Ala Arg Asp Pro Thr Ser Gly Thr Leu Tyr
35 40 45
Leu Gly Ala Thr Asn Phe Leu Phe Gln Leu Ser Pro Gly Leu Gln Leu
50 55 60
Glu Ala Thr Val Ser Thr Gly Pro Val Leu Asp Ser Arg Asp Cys Leu
65 70 75 80
Pro Pro Val Met Pro Asp Glu Cys Pro Gln Ala Gln Pro Thr Asn Asn
85 90 95
Pro Asn Gln Leu Leu Leu Val Ser Pro Gly Ala Leu Val Val Cys Gly
100 105 110
Ser Val His Gln Gly Val Cys Glu Gln Arg Arg Leu Gly Gln Leu Glu
115 120 125
Gln Leu Leu Leu Arg Pro Glu Arg Pro Gly Asp Thr Gln Tyr Val Ala
130 135 140
Ala Asn Asp Pro Ala Val Ser Thr Val Gly Leu Val Ala Gln Gly Leu
145 150 155 160
Ala Gly Glu Pro Leu Leu Phe Val Gly Arg Gly Tyr Thr Ser Arg Gly
165 170 175
Val Gly Gly Ile Pro Pro Ile Thr Thr Arg Ala Leu Trp Pro Pro
180 185 190
Asp Pro Gln Ala Ala Phe Ser Tyr Glu Glu Thr Ala Lys Leu Ala Val
195 200 205
Gly Arg Leu Ser Glu Tyr Ser His His Phe Val Ser Ala Phe Ala Arg
210 215 220
Gly Ala Ser Ala Tyr Phe Leu Phe Leu Arg Arg Asp Leu Gln Ala Gln
225 230 235 240
Ser Arg Ala Phe Arg Ala Tyr Val Ser Arg Val Cys Leu Arg Asp Gln
245 250 255
His Tyr Tyr Ser Tyr Val Glu Leu Pro Leu Ala Cys Glu Gly Arg
260 265 270

Tyr Gly Leu Ile Gln Ala Ala Ala Val Ala Thr Ser Arg Glu Val Ala
 275 280 285
 His Gly Glu Val Leu Phe Ala Ala Phe Ser Ser Ala Ala Pro Pro Thr
 290 295 300
 Val Gly Arg Pro Pro Ser Ala Ala Ala Gly Ala Ser Gly Ala Ser Ala
 305 310 315 320
 Leu Cys Ala Phe Pro Leu Asp Glu Val Asp Arg Leu Ala Asn Arg Thr
 325 330 335
 Arg Asp Ala Cys Tyr Thr Arg Glu Gly Arg Ala Glu Asp Gly Thr Glu
 340 345 350
 Val Ala Tyr Ile Glu Tyr Asp Val Asn Ser Asp Cys Ala Gln Leu Pro
 355 360 365
 Val Asp Thr Leu Asp Ala Tyr Pro Cys Gly Ser Asp His Thr Pro Ser
 370 375 380
 Pro Met Ala Ser Arg Val Pro Leu Glu Ala Thr Pro Ile Leu Glu Trp
 385 390 395 400
 Pro Gly Ile Gln Leu Thr Ala Val Ala Val Thr Met Glu Asp Gly His
 405 410 415
 Thr Ile Ala Phe Leu Gly Asp Ser Gln Gly Gln Leu His Arg Val Tyr
 420 425 430
 Leu Gly Pro Gly Ser Asp Gly His Pro Tyr Ser Thr Gln Ser Ile Gln
 435 440 445
 Gln Gly Ser Ala Val Ser Arg Asp Leu Thr Phe Asp Gly Thr Phe Glu
 450 455 460
 His Leu Tyr Val Met Thr Gln Ser Thr Leu Leu Lys Val Pro Val Ala
 465 470 475 480
 Ser Cys Ala Gln His Leu Asp Cys Ala Ser Cys Leu Ala His Arg Asp
 485 490 495
 Pro Tyr Cys Gly Trp Cys Val Leu Leu Gly Arg Cys Ser Arg Arg Ser
 500 505 510
 Glu Cys Ser Arg Gly Gln Gly Pro Glu Gln Trp Leu Trp Ser Phe Gln
 515 520 525
 Pro Glu Leu Gly Cys Leu Gln Val Ala Ala Met Ser Pro Ala Asn Ile
 530 535 540
 Ser Arg Glu Glu Thr Arg Glu Val Phe Leu Ser Val Pro Asp Leu Pro
 545 550 555 560
 Pro Leu Trp Pro Gly Glu Ser Tyr Ser Cys His Phe Gly Glu His Gln
 565 570 575
 Ser Pro Ala Leu Leu Thr Gly Ser Gly Val Met Cys Pro Ser Pro Asp
 580 585 590
 Pro Ser Glu Ala Pro Val Leu Pro Arg Gly Ala Asp Tyr Val Ser Val
 595 600 605
 Ser Val Glu Leu Arg Phe Gly Ala Val Val Ile Ala Lys Thr Ser Leu
 610 615 620
 Ser Phe Tyr Asp Cys Val Ala Val Thr Glu Leu Arg Pro Ser Ala Gln
 625 630 635 640
 Cys Gln Ala Cys Val Ser Ser Arg Trp Gly Cys Asn Trp Cys Val Trp
 645 650 655
 Gln His Leu Cys Thr His Lys Ala Ser Cys Asp Ala Gly Pro Met Val
 660 665 670
 Ala Ser His Gln Ser Pro Leu Val Ser Pro Asp Pro Pro Ala Arg Gly
 675 680 685
 Gly Pro Ser Pro Ser Pro Pro Thr Ala Pro Lys Ala Leu Ala Thr Pro
 690 695 700
 Ala Pro Asp Thr Leu Pro Val Glu Pro Gly Ala Pro Ser Thr Ala Thr

705	710	715	720
Ala Ser Asp Ile Ser Pro Gly Ala Ser Pro Ser Leu Leu Ser Pro Trp			
725	725	730	735
Gly Pro Trp Ala Gly Ser Gly Ser Ile Ser Ser Pro Gly Ser Thr Gly			
740	740	745	750
Ser Pro Leu His Glu Glu Pro Ser Pro Pro Ser Pro Gln Asn Gly Pro			
755	755	760	765
Gly Thr Ala Val Pro Ala Pro Thr Asp Phe Arg Pro Ser Ala Thr Pro			
770	770	775	780
Glu Asp Leu Leu Ala Ser Pro Leu Ser Pro Ser Glu Val Ala Ala Val			
785	785	790	795
Pro Pro Ala Asp Pro Gly Pro Glu Ala Leu His Pro Thr Val Pro Leu			
805	805	810	815
Asp Leu Pro Pro Ala Thr Val Pro Ala Thr Thr Phe Pro Gly Ala Met			
820	820	825	830
Gly Ser Val Lys Pro Ala Leu Asp Trp Leu Thr Arg Glu Gly Gly Glu			
835	835	840	845
Leu Pro Glu Ala Asp Glu Trp Thr Gly Gly Asp Ala Pro Ala Phe Ser			
850	850	855	860
Thr Ser Thr Leu Leu Ser Gly Asp Gly Asp Ser Ala Glu Leu Glu Gly			
865	865	870	875
Pro Pro Ala Pro Leu Ile Leu Pro Ser Ser Leu Asp Tyr Gln Tyr Asp			
885	885	890	895
Thr Pro Gly Leu Trp Glu Leu Glu Ala Thr Leu Gly Ala Ser Ser			
900	900	905	910
Cys Pro Cys Val Glu Ser Val Gln Gly Ser Thr Leu Met Pro Val His			
915	915	920	925
Val Glu Arg Glu Ile Arg Leu Leu Gly Arg Asn Leu His Leu Phe Gln			
930	930	935	940
Asp Gly Pro Gly Asp Asn Glu Cys Val Met Glu Leu Glu Gly Leu Glu			
945	945	950	955
Val Val Val Glu Ala Arg Val Glu Cys Glu Pro Pro Pro Asp Thr Gln			
965	965	970	975
Cys His Val Thr Cys Gln Gln His Gln Leu Ser Tyr Glu Ala Leu Gln			
980	980	985	990
Pro Glu Leu Arg Val Gly Leu Phe Leu Arg Arg Ala Gly Arg Leu Arg			
995	995	1000	1005
Val Asp Ser Ala Glu Gly Leu His Val Val Leu Tyr Asp Cys Ser Val			
1010	1010	1015	1020
Gly His Gly Asp Cys Ser Arg Cys Gln Thr Ala Met Pro Gln Tyr Gly			
1025	1025	1030	1035
Cys Val Trp Cys Glu Gly Glu Arg Pro Arg Cys Val Thr Arg Glu Ala			
1045	1045	1050	1055
Cys Gly Glu Ala Glu Ala Val Ala Thr Gln Cys Pro Ala Pro Leu Ile			
1060	1060	1065	1070
His Ser Val Glu Pro Leu Thr Gly Pro Val Asp Gly Gly Thr Arg Val			
1075	1075	1080	1085
Thr Ile Arg Gly Ser Asn Leu Gly Gln His Val Gln Asp Val Leu Gly			
1090	1090	1095	1100
Met Val Thr Val Ala Gly Val Pro Cys Ala Val Asp Ala Gln Glu Tyr			
1105	1105	1110	1115
Glu Val Ser Ser Ser Leu Val Cys Ile Thr Gly Ala Ser Gly Glu Glu			
1125	1125	1130	1135
Val Ala Gly Ala Thr Ala Val Glu Val Pro Gly Arg Gly Arg Gly Val			
1140	1140	1145	1150

Ser Glu His Asp Phe Ala Tyr Gln Asp Pro Lys Val His Ser Ile Phe
 1155 1160 1165
 Pro Ala Arg Gly Pro Arg Ala Gly Gly Thr Arg Leu Thr Leu Asn Gly
 1170 1175 1180
 Ser Lys Leu Leu Thr Gly Arg Leu Glu Asp Ile Arg Val Val Val Gly
 1185 1190 1195 1200
 Asp Gln Pro Cys His Leu Leu Pro Glu Gln Gln Ser Glu Gln Leu Arg
 1205 1210 1215
 Cys Glu Thr Ser Pro Arg Pro Thr Pro Ala Thr Leu Pro Val Ala Val
 1220 1225 1230
 Trp Phe Gly Ala Thr Glu Arg Arg Leu Gln Arg Gly Gln Phe Lys Tyr
 1235 1240 1245
 Thr Leu Asp Pro Asn Ile Thr Ser Ala Gly Pro Thr Lys Ser Phe Leu
 1250 1255 1260
 Ser Gly Gly Arg Glu Ile Cys Val Arg Gly Gln Asn Leu Asp Val Val
 1265 1270 1275 1280
 Gln Thr Pro Arg Ile Arg Val Thr Val Val Ser Arg Met Leu Gln Pro
 1285 1290 1295
 Ser Gln Gly Leu Gly Arg Arg Arg Val Val Pro Glu Thr Ala Cys
 1300 1305 1310
 Ser Leu Gly Pro Ser Cys Ser Ser Gln Gln Phe Glu Glu Pro Cys His
 1315 1320 1325
 Val Asn Ser Ser Gln Leu Ile Thr Cys Arg Thr Pro Ala Leu Pro Gly
 1330 1335 1340
 Leu Pro Glu Asp Pro Trp Val Arg Val Glu Phe Ile Leu Asp Asn Leu
 1345 1350 1355 1360
 Val Phe Asp Phe Ala Thr Leu Asn Pro Thr Pro Phe Ser Tyr Glu Ala
 1365 1370 1375
 Asp Pro Thr Leu Gln Pro Leu Asn Pro Glu Asp Pro Thr Met Pro Phe
 1380 1385 1390
 Arg His Lys Pro Gly Ser Val Phe Ser Val Glu Gly Glu Asn Leu Asp
 1395 1400 1405
 Leu Ala Met Ser Lys Glu Glu Val Val Ala Met Ile Gly Asp Gly Pro
 1410 1415 1420
 Cys Val Val Lys Thr Leu Thr Arg His His Leu Tyr Cys Glu Pro Pro
 1425 1430 1435 1440
 Val Glu Gln Pro Leu Pro Arg His His Ala Leu Arg Glu Ala Pro Asp
 1445 1450 1455
 Ser Leu Pro Glu Phe Thr Val Gln Met Gly Asn Leu Arg Phe Ser Leu
 1460 1465 1470
 Gly His Val Gln Tyr Asp Gly Glu Ser Pro Gly Ala Phe Pro Val Ala
 1475 1480 1485
 Ala Gln Val Gly Leu Gly Val Gly Thr Ser Leu Leu Ala Leu Gly Val
 1490 1495 1500
 Ile Ile Ile Val Leu Met Tyr Arg Arg Lys Ser Lys Gln Ala Leu Arg
 1505 1510 1515 1520
 Asp Tyr Lys Lys Val Gln Ile Gln Leu Glu Asn Leu Glu Ser Ser Val
 1525 1530 1535
 Arg Asp Arg Cys Lys Lys Glu Phe Thr Asp Leu Met Thr Glu Met Thr
 1540 1545 1550
 Asp Leu Thr Ser Asp Leu Leu Gly Ser Gly Ile Pro Phe Leu Asp Tyr
 1555 1560 1565
 Lys Val Tyr Ala Glu Arg Ile Phe Phe Pro Gly His Arg Glu Ser Pro
 1570 1575 1580
 Leu His Arg Asp Leu Gly Val Pro Glu Ser Arg Arg Pro Thr Val Glu

1585	1590	1595	1600
Gln Gly Leu Gly Gln	Leu Ser Asn Leu	Leu Asn Ser Lys Leu	Phe Leu
1605	1610	1615	
Thr Lys Phe Ile His	Thr Leu Glu Ser	Gln Arg Thr Phe Ser Ala Arg	
1620	1625	1630	
Asp Arg Ala Tyr Val	Ala Ser Leu	Leu Thr Val Ala	Leu His Gly Lys
1635	1640	1645	
Leu Glu Tyr Phe Thr Asp	Ile Leu Arg Thr	Leu Leu Ser Asp	Leu Val
1650	1655	1660	
Ala Gln Tyr Val Ala	Lys Asn Pro Lys	Leu Met Leu Arg Arg Thr	Glu
1665	1670	1675	1680
Thr Val Val Glu Lys	Leu Leu Thr Asn Trp	Met Ser Ile Cys	Leu Tyr
1685	1690	1695	
Thr Phe Val Arg Asp Ser Val	Gly Glu Pro	Leu Tyr Met	Leu Phe Arg
1700	1705	1710	
Gly Ile Lys His Gln Val Asp	Lys Gly Pro	Val Asp Ser Val	Thr Gly
1715	1720	1725	
Lys Ala Lys Tyr Thr Leu Asn Asp	Asn Arg Leu	Leu Arg Glu Asp	Val
1730	1735	1740	
Glu Tyr Arg Pro Leu Thr	Leu Asn Ala	Leu Ala Val Gly	Pro Gly
1745	1750	1755	1760
Ala Gly Glu Ala Gln	Gly Val Pro Val	Lys Val Leu Asp	Cys Asp Thr
1765	1770	1775	
Ile Ser Gln Ala Glu Lys	Met Leu Asp Gln	Leu Tyr Lys	Gly Val
1780	1785	1790	
Pro Leu Thr Gln Arg Pro Asp	Pro Arg Thr	Leu Asp Val	Glu Trp Arg
1795	1800	1805	
Ser Gly Val Ala Gly His	Leu Ile Leu Ser Asp	Glu Asp Val	Thr Ser
1810	1815	1820	
Glu Val Gln Gly Leu Trp	Arg Arg Leu Asn	Thr Leu Gln His	Tyr Lys
1825	1830	1835	1840
Val Pro Asp Gly Ala Thr	Val Ala Leu Val	Pro Cys Leu Thr	Lys His
1845	1850	1855	
Val Leu Arg Glu Asn Gln	Asp Tyr Val	Pro Gly Glu Arg	Thr Pro Met
1860	1865	1870	
Leu Glu Asp Val Asp	Glu Gly Ile Arg	Pro Trp His	Leu Val Lys
1875	1880	1885	
Pro Ser Asp Glu Pro Glu	Pro Pro Arg	Pro Arg Arg	Gly Ser Leu Arg
1890	1895	1900	
Gly Gly Glu Arg Glu	Arg Ala Lys	Ala Ile Pro Glu	Ile Tyr Leu Thr
1905	1910	1915	1920
Arg Leu Leu Ser Met	Lys Gly Thr	Leu Gln Lys	Phe Val Asp Asp
1925	1930	1935	Leu
Phe Gln Val Ile Leu Ser	Thr Ser Arg	Pro Val Pro	Leu Ala Val Lys
1940	1945	1950	
Tyr Phe Asp Leu Leu Asp	Glu Gln Ala Gln	Gln His Gly	Ile Ser
1955	1960	1965	
Asp Gln Asp Thr Ile His	Ile Trp Lys	Thr Asn Ser	Leu Pro Leu Arg
1970	1975	1980	
Phe Trp Ile Asn Ile Ile	Lys Asn Pro	Gln Phe Val	Phe Asp Val Gln
1985	1990	1995	2000
Thr Ser Asp Asn Met	Asp Ala Val	Leu Val Ile	Ala Gln Thr Phe
2005	2010	2015	
Met Asp Ala Cys Thr	Leu Ala Asp	His Lys Leu Gly	Arg Asp Ser Pro
2020	2025	2030	

Ile Asn Lys Leu Leu Tyr Ala Arg Asp Ile Pro Arg Tyr Lys Arg Met
 2035 2040 2045
 Val Glu Arg Tyr Tyr Ala Asp Ile Arg Gln Thr Val Pro Ala Ser Asp
 2050 2055 2060
 Gln Glu Met Asn Ser Val Leu Ala Glu Leu Ser Trp Asn Tyr Ser Gly
 2065 2070 2075 2080
 Asp Leu Gly Ala Arg Val Ala Leu His Glu Leu Tyr Lys Tyr Ile Asn
 2085 2090 2095
 Lys Tyr Tyr Asp Gln Ile Ile Thr Ala Leu Glu Glu Asp Gly Thr Ala
 2100 2105 2110
 Gln Lys Met Gln Leu Gly Tyr Arg Leu Gln Gln Ile Ala Ala Ala Val
 2115 2120 2125
 Glu Asn Lys Val Thr Asp Leu
 2130 2135

<210> 11
<211> 2190
<212> DNA
<213> HOMO SAPIEN

<400> 11

atgcctgctc	tggggccaa	gttctccag	gtctctggg	ccgggtgggt	cctcaccc	60
cagcccccc	caccaactc	attcaactcc	aatggcacgt	atctgcagca	cctggcaagg	120
gacc	caccc	ctac	ctac	cttctttcca	gctgagcc	180
gggctgc	tgagg	ccac	agtgtccacc	ggccctgtgc	tagacagcag	240
ccac	ctgt	ga	gtgccc	gcccagccta	ggactgcctg	300
ctc	ctgg	ta	gtgg	ccaacaaccc	gaatcagctg	360
gccc	agg	gg	cttgg	tgccggagcg	tgaccagg	420
cag	ggc	gg	gg	ctgctgcggc	cagagcggcc	480
caat	atgt	gg	tc	agcacggtgg	ggctggtagc	540
gcag	ggg	gg	tc	ggatacaca	ccaggggctt	600
atcc	accca	cc	gtgg	gcaagggtgt	gggggtggc	660
gagg	agac	aa	gg	ccggccgacc	ccaaagctgc	720
gcct	ttgc	cc	gg	cttccgagt	cttcccttat	780
tct	agag	ct	gtgt	acagccaca	cttcgtgagt	840
tat	gtgg	ct	gtgt	ctgtctcc	ggaccagca	900
gtgg	ccac	gt	gtgt	cttactactc	cttactactc	960
gcac	cccc	cc	gtgt	gggtgatcca	ggctgcagct	1020
ctgt	gggg	cc	gtgt	gggtgatctt	ttgcagctt	1080
tc	cc	at	gtgt	cttccctgt	cttccctgt	1140
tac	cccc	cc	gtgt	atcgacgcg	atcgacgcg	1200
aatt	tctact	gt	gtgt	atcgacgcg	atcgacgcg	1260
gtgc	ac	cc	gtgt	atcgacgcg	atcgacgcg	1320
cacac	cccc	cc	gtgt	atcgacgcg	atcgacgcg	1380
ccagg	ggatt	cc	gtgt	atcgacgcg	atcgacgcg	1440
ctgg	gtata	cc	gtgt	atcgacgcg	atcgacgcg	1500
ccat	acttcc	cc	gtgt	atcgacgcg	atcgacgcg	1560
gggac	cac	cc	gtgt	atcgacgcg	atcgacgcg	1620
ccct	cc	cc	gtgt	atcgacgcg	atcgacgcg	1680
ctgact	gtgt	cc	gtgt	atcgacgcg	atcgacgcg	1740
agagg	ggcc	cc	gtgt	atcgacgcg	atcgacgcg	1800
aaaacttcc	tcttttctt	cc	gtgt	atcgacgcg	atcgacgcg	1860
tgccagg	gtgt	cc	gtgt	atcgacgcg	atcgacgcg	1920
acccaca	agg	cc	gtgt	atcgacgcg	atcgacgcg	1980
		tgctgg	gtgt	atcgacgcg	atcgacgcg	2040

cagcagagct tgagggccct cccgcccccc tcatccccc gtcgcgcctc gactaccagt	2100
atgacacccc cgggctctgg gagcttggaa aggcgcacctt gggggcaagc tcctgcccct	2160
gtgtggagag cgttcaggc tccacgttga	2190

<210> 12
<211> 729
<212> PRT
<213> HOMO SAPIEN

<400> 12	
Met Pro Ala Leu Gly Pro Ala Leu Leu Gln Ala Leu Trp Ala Gly Trp	
1 5 10 15	
Val Leu Thr Leu Gln Pro Leu Pro Pro Thr Ala Phe Thr Pro Asn Gly	
20 25 30	
Thr Tyr Leu Gln His Leu Ala Arg Asp Pro Thr Ser Gly Thr Leu Tyr	
35 40 45	
Leu Gly Ala Thr Asn Phe Leu Phe Gln Leu Ser Pro Gly Leu Gln Leu	
50 55 60	
Glu Ala Thr Val Ser Thr Gly Pro Val Leu Asp Ser Arg Asp Cys Leu	
65 70 75 80	
Pro Pro Val Met Pro Asp Glu Cys Pro Gln Ala Gln Pro Thr Asn Asn	
85 90 95	
Pro Asn Gln Leu Leu Leu Val Ser Pro Gly Ala Leu Val Val Cys Gly	
100 105 110	
Ser Val His Gln Gly Val Cys Glu Gln Arg Arg Leu Gly Gln Leu Glu	
115 120 125	
Gln Leu Leu Leu Arg Pro Glu Arg Pro Gly Asp Thr Gln Tyr Val Ala	
130 135 140	
Ala Asn Asp Pro Ala Val Ser Thr Val Gly Leu Val Ala Gln Gly Leu	
145 150 155 160	
Ala Gly Glu Pro Leu Leu Phe Val Gly Arg Gly Tyr Thr Ser Arg Gly	
165 170 175	
Val Gly Gly Gly Ile Pro Pro Ile Thr Thr Arg Ala Leu Trp Pro Pro	
180 185 190	
Asp Pro Gln Ala Ala Phe Ser Tyr Glu Glu Thr Ala Lys Leu Ala Val	
195 200 205	
Gly Arg Leu Ser Glu Tyr Ser His His Phe Val Ser Ala Phe Ala Arg	
210 215 220	
Gly Ala Ser Ala Tyr Phe Leu Phe Leu Arg Arg Asp Leu Gln Ala Gln	
225 230 235 240	
Ser Arg Ala Phe Arg Ala Tyr Val Ser Arg Val Cys Leu Arg Asp Gln	
245 250 255	
His Tyr Tyr Ser Tyr Val Glu Leu Pro Leu Ala Cys Glu Gly Gly Arg	
260 265 270	
Tyr Gly Leu Ile Gln Ala Ala Ala Val Ala Thr Ser Arg Glu Val Ala	
275 280 285	
His Gly Glu Val Leu Phe Ala Ala Phe Ser Ser Ala Ala Pro Pro Thr	
290 295 300	
Val Gly Arg Pro Pro Ser Ala Ala Ala Gly Ala Ser Gly Ala Ser Ala	
305 310 315 320	
Leu Cys Ala Phe Pro Leu Asp Glu Val Asp Arg Leu Ala Asn Arg Thr	
325 330 335	
Arg Asp Ala Cys Tyr Thr Arg Glu Gly Arg Ala Glu Asp Gly Thr Glu	
340 345 350	
Val Ala Tyr Ile Glu Tyr Asp Val Asn Ser Asp Cys Ala Gln Leu Pro	

355	360	365
Val Asp Thr Leu Asp Ala Tyr Pro Cys Gly Ser Asp His Thr Pro Ser		
370	375	380
Pro Met Ala Ser Arg Val Pro Leu Glu Ala Thr Pro Ile Leu Glu Trp		
385	390	395
400		
Pro Gly Ile Gln Leu Thr Ala Val Ala Val Thr Met Glu Asp Gly His		
405	410	415
Thr Ile Ala Phe Leu Gly Asp Ser Gln Gly Gln Leu His Arg Val Tyr		
420	425	430
Leu Gly Pro Gly Ser Asp Gly His Pro Tyr Ser Thr Gln Ser Ile Gln		
435	440	445
Gln Gly Ser Ala Val Ser Arg Asp Leu Thr Phe Asp Gly Thr Phe Glu		
450	455	460
His Leu Tyr Val Met Thr Gln Ser Thr Leu Leu Lys Val Pro Val Ala		
465	470	475
480		
Ser Cys Ala Gln His Leu Asp Cys Ala Ser Cys Leu Ala His Arg Asp		
485	490	495
Pro Tyr Cys Gly Trp Cys Val Leu Leu Gly Arg Cys Ser Arg Arg Ser		
500	505	510
Glu Cys Ser Arg Gly Gln Gly Pro Glu Gln Trp Leu Trp Ser Phe Gln		
515	520	525
Pro Glu Leu Gly Cys Leu Gln Val Ala Ala Met Ser Pro Ala Asn Ile		
530	535	540
Ser Arg Glu Glu Thr Arg Glu Val Phe Leu Ser Val Pro Asp Leu Pro		
545	550	555
560		
Pro Leu Trp Pro Gly Glu Ser Tyr Ser Cys His Phe Gly Glu His Gln		
565	570	575
Ser Pro Ala Leu Leu Thr Gly Ser Gly Val Met Cys Pro Ser Pro Asp		
580	585	590
Pro Ser Glu Ala Pro Val Leu Pro Arg Gly Ala Asp Tyr Val Ser Val		
595	600	605
Ser Val Glu Leu Arg Phe Gly Ala Val Val Ile Ala Lys Thr Ser Leu		
610	615	620
Ser Phe Tyr Asp Cys Val Ala Val Thr Glu Leu Arg Pro Ser Ala Gln		
625	630	635
640		
Cys Gln Ala Cys Val Ser Ser Arg Trp Gly Cys Asn Trp Cys Val Trp		
645	650	655
Gln His Leu Cys Thr His Lys Ala Ser Cys Asp Ala Gly Pro Met Val		
660	665	670
Ala Ser His Gln Val Met Glu Thr Gln Gln Ser Leu Arg Ala Leu Pro		
675	680	685
Pro Pro Ser Ser Ser Arg Pro Ala Ser Thr Thr Ser Met Thr Pro Pro		
690	695	700
Gly Ser Gly Ser Trp Lys Arg Arg Pro Trp Gly Gln Ala Pro Ala Pro		
705	710	715
720		
Val Trp Arg Ala Phe Arg Ala Pro Arg		
725		

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
1 March 2001 (01.03.2001)

PCT

(10) International Publication Number
WO 01/14420 A3

(51) International Patent Classification⁷: C12N 15/12,
15/62, 15/63, C07K 14/705, 16/28, C12P 21/02, A61K
38/17, 39/395, G01N 33/53

the University of California, 12th floor, 1111 Franklin
Street, Oakland, CA 94607-5200 (US). **TAMAGNONE,**
Luca [IT/IT]; Corso Einaudi, 43, I-10129 Torino (IT).

(21) International Application Number: PCT/US00/23365

(74) Agent: **COX, Niki, D.**: Biogen, Inc., 14 Cambridge Center,
Cambridge, MA 02142 (US).

(22) International Filing Date: 25 August 2000 (25.08.2000)

(81) Designated States (national): AE. AG. AL. AM. AT. AU.
AZ. BA. BB. BG. BR. BY. BZ. CA. CH. CN. CR. CU. CZ.
DE. DK. DM. DZ. EE. ES. FI. GB. GD. GE. GH. GM. HR.
HU. ID. IL. IN. IS. JP. KE. KG. KP. KR. KZ. LC. LK. LR.
LS. LT. LU. LV. MA. MD. MG. MK. MN. MW. MX. MZ.
NO. NZ. PL. PT. RO. RU. SD. SE. SG. SI. SK. SL. TJ. TM.
TR. TT. TZ. UA. UG. US. UZ. VN. YU. ZA. ZW.

(25) Filing Language: English

(84) Designated States (regional): ARIPO patent (GH. GM.
KE. LS. MW. MZ. SD. SL. SZ. TZ. UG. ZW). Eurasian
patent (AM. AZ. BY. KG. KZ. MD. RU. TJ. TM). European
patent (AT. BE. CH. CY. DE. DK. ES. FI. FR. GB. GR. IE.
IT. LU. MC. NL. PT. SE). OAPI patent (BF. BJ. CF. CG.
CI. CM. GA. GN. GW. ML. MR. NE. SN. TD. TG).

(26) Publication Language: English

(30) Priority Data:
60/150,576 25 August 1999 (25.08.1999) US

Published:

— with international search report

(71) Applicants (for all designated States except US): **UNIVERSITY OF TORINO** [IT/IT]; Department of Biomedical Sciences and Human Oncology, IRCC, SP 142, I-10060 Candiolo (IT). **REGENTS OF THE UNIVERSITY OF CALIFORNIA** [US/US]; 12th floor, 1111 Franklin Street, Oakland, CA 94607-5200 (US).

(88) Date of publication of the international search report:
10 May 2002

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(72) Inventors; and

(75) Inventors/Applicants (for US only): **ARTIGIANI, Stefania** [IT/IT]; Corso Brunelleschi, 121/B, I-10100 Torino (IT). **COMOGLIO, Paolo, M.** [IT/IT]; Strada Valsalice, 183/8, I-10100 Torino (IT). **GOODMAN, Corey, S.** [US/US]; Regents of the University of California, 12th floor, 1111 Franklin Street, Oakland, CA 94607-5200 (US). **TESIER-LAVIGNE, Marc** [US/US]; Regents of

WO 01/14420 A3

(54) Title: NOVEL MEMBERS OF THE PLEXIN FAMILY AND USES THEREOF

(57) Abstract: The invention provides methods and compositions related to novel plexins. The polypeptides may be produced recombinantly from transformed host cells and from the disclosed plexin encoding nucleic acids or purified from human cells. The invention provides isolated plexin hybridization probes and primers capable of specifically hybridizing with the disclosed plexin genes, plexin-specific binding agents such as specific antibodies, and methods of making and using the subject compositions in diagnosis, therapy and in biopharmaceutical industry. The invention also provides novel plexin neuropilin multimeric receptor complexes for semaphorins and methods of use thereof, including but not limited to, the treatment and diagnosis of neurological disease and neuroregeneration, immune modulation, and viral and oncological diseases.

II INTERNATIONAL SEARCH REPORT

International Application No

PCT/US 00/23365

A. CLASSIFICATION OF SUBJECT MATTER

IPC 7	C12N15/12	C12N15/62	C12N15/63	C07K14/705	C07K16/28
	C12P21/02	A61K38/17	A61K39/395	G01N33/53	

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, WPI Data, PAJ, BIOSIS, MEDLINE, EMBASE, SCISEARCH, STRAND

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	EMBL DATABASE EMHUM4:HSAB2313; ACCESSION-NO: AB002313, 1 July 1997 (1997-07-01), XP002157964 the whole document	1-5
Y	& DATABASE SWALL:015031; ACCESSION-NO: 015031, 1 January 1998 (1998-01-01), the whole document & NAGASE, T. ET AL.: "Prediction of the coding sequences of unidentified human genes. VII. The complete sequences of 100 new cDNA clones from brain which can code for large proteins in vitro" DNA RESEARCH, vol. 4, 1997, pages 141-150, XP002102085 page 142 -page 150 'Results and Discussion' figure 3; tables 1,2 ----- -/-	6-9

 Further documents are listed in the continuation of box C. Patent family members are listed in annex.

° Special categories of cited documents :

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search

Date of mailing of the international search report

6 August 2001

31.08.01

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl.
Fax: (+31-70) 340-3016

Authorized officer

Donath, C

II INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 00/23365

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	EMBL DATABASE EM_HUM:AB014520; ACCESSION-NO.:AB014520, 15 July 1998 (1998-07-15), XP002173834 the whole document	1-5
Y	& ISHIKAWA, K.-I. ET AL.: "Prediction of the coding sequences of unidentified human genes. X. The complete sequences of 100 new cDNA clones from brain which can code for large proteins in vitro" DNA RESEARCH, vol. 5, 30 June 1998 (1998-06-30), pages 169-176, XP002121149 page 172 -page 176 * 'Results and Discussion' * figures 1,2; tables 1-3 ---	6-9
X	EMBL DATABASE EMHUM4:HS5211110; ACCESSION-NO: U52111, 9 May 1996 (1996-05-09), XP002173835 page 12 -page 13 * Gene="PLXB3" and product="plexin-related protein" *	1-5
Y	---	6-9
X	EMBL DATABASE EMHUM6:HSOCTPROT; ACCESSION-NO: X87831, 6 February 1996 (1996-02-06), XP002173836 the whole document	1-3
X	EMBL DATABASE EM_OV:XLPLEX; ACCESSION-NO:D38175, 25 August 1995 (1995-08-25), XP002173837 the whole document & OHTA, K. ET AL.: "Plexin: a novel neuronal cell surface molecule that mediates cell adhesion via a homophilic binding mechanism in the presence of calcium ions" NEURON, vol. 14, 1995, pages 1189-1199, XP001013227 ---	1-3
X	WO 99 04263 A (THE JOHN HOPKINS UNIVERSITY SCHOOL OF MEDICINE) 28 January 1999 (1999-01-28)	10,11
Y	page 5, line 9 -page 10, line 16 page 16, line 6 -page 22, line 3 page 23, line 22 -page 24, line 18 page 29, line 6 -page 32, line 4 ---	12,13
	-/--	

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 00/23365

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	KAMEYAMA, T. ET AL.: "Identification of a cell surface protein plexin (the B2) in mouse, and its expression in developing nervous systems" NEUROSCIENCE RESEARCH SUPPLEMENT, vol. 18, 1993, page S115 XP000945106 the whole document ---	6-9
Y	COMEAU, M. ET AL.: "A poxvirus-encoded semaphorin induces cytokine production from monocytes and binds to a novel cellular semaphorin receptor, VESPR" IMMUNITY, vol. 8, April 1998 (1998-04), pages 473-482, XP000945259 cited in the application page 478 -page 480 'Discussion' ---	12,13
A	MAESTRINI, E. ET AL.: "A family of transmembrane proteins with homology to the MET-hepatocyte growth factor receptor" PROC.NATL.ACAD.SCI.USA, vol. 93, no. 2, 1996, pages 674-678, XP000941746 the whole document ---	1-9
P,X	TAMAGNONE, L. ET AL.: "Plexins are a large family of receptors for transmembrane, secreted, and GPI-anchored semaphorins in vertebrates" CELL, vol. 99, 1 October 1999 (1999-10-01), pages 71-80, XP000941702 page 72 -page 78 'Results' and 'Discussion' ---	1-5
P,Y	TAKAHASHI, T. ET AL.: "Plexin-Neuropilin-1 complexes form functional semaphorin-3A receptors" CELL, vol. 99, 1 October 1999 (1999-10-01), pages 59-69, XP000941701 page 60 -page 67 'Results' and 'Discussion' ---	6-9,12, 13
A	NAKAMURA, F. ET AL.: "Neuropilin-1 extracellular domains mediate semaphorin D/III-induced growth cone collapse" NEURON, vol. 21, November 1998 (1998-11), pages 1093-1100, XP002174004 cited in the application the whole document ---	10-13

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 00/23365

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. Claims Nos.: because they relate to subject matter not required to be searched by this Authority, namely:
see FURTHER INFORMATION sheet PCT/ISA/210
2. Claims Nos.: 10,11 (partially)
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
see FURTHER INFORMATION sheet PCT/ISA/210
3. Claims Nos.: because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
2. As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:
4. No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

The additional search fees were accompanied by the applicant's protest.
 No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/SA/ 210

Continuation of Box I.1

Although claims 10-13 are directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.

Although claim 14 is directed to a diagnostic method practised on the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.

Continuation of Box I.2

Claims Nos.: 10,11 (partially)

Claims 10 and 11 concern methods which comprise the administration of an agent capable of interfering with the association between a plexin and a neuropilin. Since in the specification this agent is exemplified only to be an antibody raised against the plexin and since it is completely unclear which kind of substances besides said antibody also will be capable of interfering with the association between a plexin and a neuropilin, the scope of said claims is totally ambiguous and undefined as far as any kind of substance other than an antibody raised against the plexin is concerned.

Therefore, the search in respect of claims 10 nad 11 has been limited to methods comprising the administration of an antibody raised against the plexin and which is capable of interfering with the assiciation between a plexin and a neuropilin.

The applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. Claims: 1-9,12-14 (partially)

Claims 1-9,12-14 (partially) refer to the isolation and cloning of a member of the plexin family, plexin B-2. Antibodies specifically binding to this polypeptide, a fusion protein, and methods of diagnosing for tumors, treating a disorder involving aberrant immune regulation involving a signal pathway between plexin and a neuropilin, and suppressing aberrant cell growth, all methods by using either the polypeptide or the antibodies directed against the plexin B-2.

2. Claims: 1-9,12-14 (partially)

Claims 1-9,12-14 (partially) refer to the isolation and cloning of a member of the plexin family, plexin B-3. Antibodies specifically binding to this polypeptide, a fusion protein, and methods of diagnosing for tumors, treating a disorder involving aberrant immune regulation involving a signal pathway between plexin and a neuropilin, and suppressing aberrant cell growth, all methods by using either the polypeptide or the antibodies directed against the plexin B-3.

3. Claims: 1-9,12-14 (partially)

Claims 1-9,12-14 (partially) refer to the isolation and cloning of a member of the plexin family, plexin D-1. Antibodies specifically binding to this polypeptide, a fusion protein, and methods of diagnosing for tumors, treating a disorder involving aberrant immune regulation involving a signal pathway between plexin and a neuropilin, and suppressing aberrant cell growth, all methods by using either the polypeptide or the antibodies directed against the plexin D-1.

4. Claims: 1-9,12-14 (partially)

Claims 1-9,12-14 (partially) refer to the isolation and cloning of a member of the plexin family, plexin A-4. Antibodies specifically binding to this polypeptide, a fusion protein, and methods of diagnosing for tumors, treating a disorder involving aberrant immune regulation involving a signal pathway between plexin and a neuropilin, and suppressing aberrant cell growth, all methods by using either the polypeptide or the antibodies directed against the plexin A-4.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

5. Claims: 10,11

Claims 10 and 11 refer either to a method of suppressing or altering aberrant cell growth involving a signalling pathway between a plexin and a neuropilin in a mammal or to a method of treating, suppressing or altering a disorder involving aberrant immune regulation involving a signalling pathway between a plexin and a neuropilin in a mammal, both methods comprises the administration of an agent in general being capable of interfering with the association between the plexin and neuropilin to said mammal.

INTERNATIONAL SEARCH REPORT

Information on patent family members

Int. dional Application No

PCT/US 00/23365

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9904263 A	28-01-1999	AU 8405398 A	10-02-1999